

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.

THIS PAGE BLANK (USPTO)

BEST AVAILABLE COPY

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 6, 2003, 11:16:00 : Search time 27.8333 Seconds
(without alignments)
19.150 Million cell updates/sec

Title: PAT943-1

Perfect score: 23

Sequence: 1 dyda 4

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 710

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_101002:*

- 1: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
- 2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
- 3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
- 4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
- 5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
- 6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
- 7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
- 8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
- 9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
- 10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
- 11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
- 12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
- 13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
- 14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
- 15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
- 16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
- 17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
- 18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
- 19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
- 20: /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
- 21: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
- 22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
- 23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	23	100.0	13	20	Thermus thermophil
2	23	100.0	16	19	CDR-2 of the H cha
3	23	100.0	19	21	YadA homologous pe
4	23	100.0	20	17	Alpha-1,3/4-fucosi
5	23	100.0	21	21	Proteobacterial ex
6	23	100.0	23	18	Cyclitol-ubiquinol
7	23	100.0	28	15	Sequence of trypti
8	23	100.0	38	22	Peptide #10124 enc
9	23	100.0	38	22	Human brain expres
10	23	100.0	38	22	Human bone marrow

11	23	100.0	38	22	AA36432	Peptide #10469 enc
12	23	100.0	41	21	AAG10870	Arabidopsis thalia
13	23	100.0	41	21	AAG54078	Arabidopsis thalia
14	23	100.0	44	19	AAW79339	Staphylococcus aur
15	23	100.0	45	19	AAW62275	Streptococcus pneu
16	23	100.0	45	20	AA41839	S. pneumoniae trun
17	23	100.0	45	20	AAW92457	Partial D. gossyp
18	23	100.0	49	17	AAW04953	Human apoptosis re
19	23	100.0	58	22	ABB11106	Propionibacterium
20	23	100.0	61	22	AAU67058	Peptide having aff
21	23	100.0	64	19	AAW62591	Duffy blood group
22	23	100.0	66	16	AAW73956	Peptide having aff
23	23	100.0	66	19	AAW62590	Homo sapiens anti-
24	23	100.0	66	19	AAW54861	Clostridium thermo
25	23	100.0	67	21	AA12795	Novel human diagno
26	23	100.0	68	22	ABG23865	Peptide #10253 enc
27	23	100.0	69	22	ABB42747	Protein #8045 enco
28	23	100.0	69	22	AAW63638	Human brain expres
29	23	100.0	69	22	AAW76452	Human bone marrow
30	23	100.0	69	22	AAW6452	Peptide #10598 enc
31	23	100.0	69	22	AAW36561	Human peptide enco
32	23	100.0	69	23	ABG45716	Sequence encoded b
33	23	100.0	70	9	AAW81871	Arabidopsis thalia
34	23	100.0	70	21	AAG10868	Arabidopsis thalia
35	23	100.0	70	21	AAG54077	Human testicular a
36	23	100.0	79	22	ABW95902	Human reproductive
37	23	100.0	79	22	AAW95203	Staphylococcus epi
38	23	100.0	79	23	ABP38575	H. pylori transpor
39	23	100.0	81	18	AAW20370	Arabidopsis thalia
40	23	100.0	81	21	AAG19542	Novel human diagno
41	23	100.0	81	22	ABG26683	Subtilase SEEP f
42	23	100.0	84	19	AAW62245	Human ORFX protein
43	23	100.0	86	23	ABP05712	Human ovarian anti
44	23	100.0	87	23	ABP42351	Arabidopsis thalia
45	23	100.0	91	21	AAW61460	

ALIGNMENTS

RESULT 1

AA50026

ID AA50026 standard; peptide; 13 AA.

XX AA50026;

AC AA50026;

XX AA50026;

DT 19-JAN-2000 (first entry)

XX Thermus thermophilus Pol III-type enzyme alpha subunit peptide TTH2.

DE DNA polymerase III; dnaE gene; alpha subunit; thermostable;

XX DNA synthesis; speed; accuracy; processivity; frameshift; holoenzyme;

KW DNA assembly; exonuclease; proofreading; inverse PCR.

KW Thermus thermophilus.

OS OS

XX WO9953074-A1.

PN 21-OCT-1999.

XX 09-APR-1998; 98WO-US07070.

XX 09-APR-1998; 98WO-US07070.

XX (UYRQ) UNIV ROCKEFELLER.

XX Yurieva O, Kuriyan J, O'Donnell ME, Jeruzalmi D;

XX WPI; 1999-611306/52.

XX New isolated thermostable DNA polymerase III-type enzyme, used

PT particularly for the amplification and sequencing of nucleic acids

XX

PS Example 10; Fig 15B; 156pp; English.
 XX
 CC This sequence represents TTH2 peptide, derived from the purified
 CC alpha subunit (encoded by the dnaE gene) of a novel *Thermus thermophilus*
 CC thermostable DNA polymerase holoenzyme which corresponds to DNA
 CC polymerase III (Pol III) of *E. coli*. The peptide was used to design
 CC primers (AA230923-230924) for PCR of the dnaE gene. Pol III-type
 CC enzymes have a high processivity (>50 kb) and rapid rate of synthesis
 CC (750 nucleotides/s). Pol III consists of 18 subunits of 10 different
 CC types. The DNA polymerase core is the catalytic unit and consists of the
 CC alpha (DNA polymerase), epsilon (3'-5' exonuclease) and theta subunits.
 CC The beta subunit ("sliding clamp") is ring-shaped and encircles DNA and
 CC slides along it while tethering the Pol III holoenzyme to the template.
 CC It is the beta subunit which is responsible for the high processivity
 CC and speed. The gamma complex is responsible for the high processivity
 CC psi subunits) is the "clamp loader" composed of gamma, delta, chi and
 CC assembly of beta clamps around DNA. A dimer of the tau subunit acts as a
 CC "macromolecular organizer", holding together molecules of core
 CC polymerase and one molecule of gamma complex, forming the Pol III*
 CC subassembly. Two beta dimers associate with the two cores within Pol
 CC III* to form the holoenzymes capable of replicating both strands of
 CC duplex DNA simultaneously. The *Thermus thermophilus* Pol III-type enzyme
 CC can be used in molecular cloning techniques such as PCR (polymerase
 CC chain reaction). The current limitations of enzymes previously used in
 CC PCR are that they are unable to synthesise extended lengths of
 CC nucleotides, and in the instance of Taq (*Thermus aquaticus*) polymerase,
 CC the lack of 3'-5' exonuclease activity and the subsequent inability to
 CC excise misinserted nucleotides (proofreading).
 XX
 SQ Sequence 13 AA;
 Query Match 100.0%; Score 23; DB 20; Length 13;
 Best Local Similarity 100.0%; Pred. No. 57;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DYDA 4
 Db | | | |
 5 DYDA 8
 RESULT 2
 AAW58532
 ID AAW58532 standard; peptide; 16 AA.
 AC AAW58532;
 XX
 XX
 DT 02-SEP-1998 (first entry)
 DE CDR-2 of the H chain variable region from mouse anti-My-10.
 XX
 KW Mouse; anti-My-10; H chain variable region; L chain variable region;
 KW antibody; heavy; light; CDR; human CD34 antigen; hybridoma; anti-CD34.
 XX
 OS Mus sp.
 XX
 PN JP10155489-A.
 XX
 PD 16-JUN-1998.
 XX
 PF 27-NOV-1996; 96JP-0331647.
 XX
 PR 27-NOV-1996; 96JP-0331647.
 XX
 PA (ASAH) ASAH KASEI KOGYO KK.
 PA (ASAH) ASAH MEDICAL CO LTD.
 XX
 XX WPI; 1998-391043/34.
 XX
 PT Recombinant antibody against human CD34 - and nucleic acid encoding
 PT it, used for efficient production of the antibody
 XX
 PS Claim 1; Page 2; 15pp; Japanese.
 XX

CC The present invention describes a nucleic acid which encodes an
 CC anti-My10 antibody. The present sequence represents the CDR-2 of
 CC the H chain variable region of the antibody. Also described is a
 CC method for the production of a recombinant antibody in which the above
 CC nucleic acid is used to produce an antibody which combines with human
 CC CD34 antigen by gene recombination. The anti-CD34 antibody gene is used
 CC to produce recombinant anti-CD34 antibodies efficiently, which can be
 CC used in pharmaceuticals.
 XX
 SQ Sequence 16 AA;
 Query Match 100.0%; Score 23; DB 19; Length 16;
 Best Local Similarity 100.0%; Pred. No. 72;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DYDA 4
 Db | | | |
 9 DYDA 12
 RESULT 3
 AAB51592
 ID AAB51592 standard; Peptide; 19 AA.
 AC AAB51592;
 XX
 DT 15-FEB-2001 (first entry)
 XX
 DE Yada homologous peptide #55.
 XX
 KW Proteobacteria; extracellular domain; virulence determinant; Yada;
 KW adhesin; proteobacterial infection prevention; vaccine.
 XX
 OS *Yersinia pestis*.
 XX
 PN WO2000061165-A1.
 XX
 PD 19-OCT-2000.
 XX
 PF 13-APR-2000; 2000WO-US09866.
 XX
 PR 13-APR-1999; 99US-0129073.
 XX
 PA (SMIK) SMITHKLINE BEECHAM CORP.
 PA (SMIK) SMITHKLINE BEECHAM PLC.
 XX
 PI Lupas AN;
 XX
 DR WPI; 2000-647397/62.
 XX
 PT An isolated polypeptide conserved in proteobacterial extracellular
 PT domains used in the treatment and prevention of bacterial infections -
 XX
 PS Example 5; Page 60; 85pp; English.
 XX
 CC This invention relates to peptides AAB51512 - AAB51537 which represent
 CC conserved proteobacterial extracellular domains. Sequences
 CC AAB51538 - AAB51618 represent peptides homologous to Yada, a *Yersinia*
 CC adhesin which is an important virulence determinant of the *Yersinia*
 CC species. The invention includes an antibody which binds to the
 CC proteobacterial extracellular peptides, and an immunogenic composition
 CC containing the antibody used as a vaccine to prevent infection by a
 CC proteobacteria. The polypeptides and antibodies are useful in the
 CC treatment and prevention of proteobacterial infections. The polypeptides
 CC can also be used to identify compounds which antagonize the binding of a
 CC bacterial adhesion to its ligand. The host cell can be used to produce
 CC the polypeptides in a suitable culture system. The composition can be
 CC used to vaccinate a patient against a proteobacterial infection.
 XX
 SQ Sequence 19 AA;
 Query Match 100.0%; Score 23; DB 21; Length 19;
 Best Local Similarity 100.0%; Pred. No. 86;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DYDA 4
Db 10 DYDA 13

RESULT 4
AAR86991
ID AAR86991 standard; peptide; 20 AA.
XX AC AAR86991;
XX DT 05-JUL-1996 (first entry)
XX DE Alpha-1,3/4-fucosidase peptidic fragment 57.
XX KW Streptomyces; alpha-1,3/4-fucosidase; primer; PCR; amplification; probe;
KW host organism; E.coli; animal; plant; structure; function; sugar chain;
KW glycoprotein.
XX OS Streptomyces sp. 142.
XX PN JP07308192-A.
XX PD 28-NOV-1995.
XX PF 18-MAY-1994; 94JP-0127109.
XX PR 18-MAY-1994; 94JP-0127109.
XX PA (TAKI) TAKARA SHUZO CO LTD.
XX PI Ikunoshin K, Masanori M, Mutsumi S, Yuka T;
XX WPI; 1996-043973/05.
XX PT Alpha-1,3/4-fucosidase gene from Streptomyces sp. 142 - useful for
XX recombinant production of the enzyme
XX PS Disclosure; Page 13; 16pp; Japanese.
XX CC The peptide fragments AAR86990-96 are partial amino acid sequences
CC derived from the purified alpha-1,3/4-fucosidase from Streptomyces
CC sp.142. The fragments were used to generate the primers AAT08084-97 used
CC for cloning the full length gene encoding the alpha-1,3/4-fucosidase.
CC The gene can be used to transform host organisms, e.g. E.coli, animal or
CC plants cells, to produce alpha-1,3/4-fucosidase. The protein can be used
CC for determining the structure and function of sugar chains or
XX glycoproteins.

SQ Sequence 20 AA;
Query Match 100.0%; Score 23; DB 17; Length 20;
Best Local Similarity 100.0%; Pred. No. 91;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
Db 10 DYDA 13
RESULT 5
AAB51521
ID AAB51521 standard; Peptide; 21 AA.
XX AC AAB51521;
XX DT 15-FEB-2001 (first entry)
XX DE Proteobacterial extracellular conserved peptide SEQ ID 10.
XX KW Proteobacteria; extracellular domain; virulence determinant; Yada;

KW adhesin; proteobacterial infection prevention; vaccine.
XX Yersinia pestis.
XX WO200061165-A1.
XX PD 19-OCT-2000.
XX PF 13-APR-2000; 2000WO-US09866.
XX PR 13-APR-1999; 99US-0129073.
XX PA (SMIK) SMITHKLINE BEECHAM CORP.
XX PA (SMIK) SMITHKLINE BEECHAM PLC.
XX PI Lupas AN;
XX WPI; 2000-647397/62.
XX PT An isolated polypeptide conserved in proteobacterial extracellular
XX domains used in the treatment and prevention of bacterial infections -
XX Claim 3; Page 64; 85pp; English.
XX CC This invention relates to peptides AAB51512 - AAB51537 which represent
CC conserved proteobacterial extracellular domains. Sequences
CC AAB51538 - AAB51618 represent peptides homologous to yada, a yersinia
CC adhesin which is an important virulence determinant of the Yersinia
CC species. The invention includes an antibody which binds to the
CC proteobacterial extracellular peptides, and an immunogenic composition
CC containing the antibody used as a vaccine to prevent infection by a
CC proteobacteria. The polypeptides and antibodies are useful in the
CC treatment and prevention of proteobacterial infections. The polypeptides
CC can also be used to identify compounds which antagonize the binding of a
CC bacterial adhesion to its ligand. The host cell can be used to produce
CC the polypeptides in a suitable culture system. The composition can be
CC used to vaccinate a patient against a proteobacterial infection.
XX SQ Sequence 21 AA;

Query Match 100.0%; Score 23; DB 21; Length 21;
Best Local Similarity 100.0%; Pred. No. 96;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DYDA 4
Db 12 DYDA 15
RESULT 6
AAW36471
ID AAW36471 standard; Peptide; 23 AA.
XX AC AAW36471;
XX DT 25-FEB-1998 (first entry)
XX DE Cyclitol-ubiquinone-oxidoreductase peptide 14.
XX KW oxidation; cyclitol; cyclitol-ubiquinone oxidoreductase; synthesis;
XX ketocyclitol; inosose epimer.
XX OS Gluconobacter oxidans DSM 50049.
XX FH Key Location/Qualifiers
FT Misc-difference 6 /label= unknown
FT /note= "encoded by NNN"
XX PN DE19628873-A1.
XX PD 23-JAN-1997.
XX

PF 17-JUL-1996; 96DE-1028873.
 XX
 PR 17-JUL-1995; 95DE-1025990.
 XX
 PA (FRAU) FRAUNHOFER GES FOERDERUNG ANGEWANDTEN.
 XX
 PI Freivogel KW, Wiesner W, Wissler JH;
 XX
 DR WPI; 1997-088583/09.
 XX
 DR N-PSDB; AAT96623.
 XX
 PT Cyclitol-ubiquinone oxido:reductase enzyme - useful for analysis and
 PT synthesis of keto:cyclitol epimers, etc.
 XX
 PS Claim 5; Page 15; 16pp; German.
 XX
 CC The peptide sequences shown in AAW36454-74 are fragments of a novel
 CC enzyme capable of oxidising cyclitols and their derivatives. The enzyme,
 CC a cyclitol-ubiquinone oxidoreductase, has a molecular weight of 87 kD.
 CC DNA sequences derived by reverse translation of the partial peptide
 CC sequences were also generated (see AAT96607-26). The enzyme is useful as
 CC a catalyst for the analysis and synthesis of ketocyclitol epimers,
 CC especially inosose epimers, preferably using coenzyme Q. It is also
 CC useful as an antigen for antibody production or analysis and for use in
 CC immunossays. The corresponding cDNA is useful for the analysis and
 CC synthesis of molecular-biological equivalence structures and associated
 CC functions.
 XX
 XX
 SQ Sequence 23 AA;
 Query Match 100.0%; Score 23; DB 18; Length 23;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DYDA 4
 Db ||||
 17 DYDA 20
 RESULT 7
 AAR49685
 ID AAR49685 standard; Peptide: 28 AA.
 XX
 AC AAR49685;
 XX
 DT 18-AUG-1994 (first entry)
 XX
 DE Sequence of tryptic peptide T4 (P4) of sweet corn starch primer
 DE protein designated amylogenin.
 XX
 KW Starch primer; amylogenin; starch synthesis; SGP; maize;
 KW sweet corn; self glucosylation protein; tryptic peptide.
 XX
 OS Zea mays.
 XX
 PN WO9404693-A.
 XX
 XX 03-MAR-1994.
 PD
 XX 26-AUG-1993; 93WO-GB01821.
 PF
 XX 26-AUG-1992; 92GB-0018185.
 PR
 XX (ZENE) ZENECA LTD.
 PA
 XX Gieowar-singh D, Keeling PL, Lomako J, Singletary GW;
 PI Whelan WJ;
 PI WPI; 1994-083207/10.
 DR
 XX Plants, esp. maize, having altered starch production ability -
 PT prepared by inserting into the plant genome an amylogenin gene
 PT encoding a starch primer which initiates synthesis.

XX
 PS Example; page 35; 85pp; English.
 XX
 CC Amylogenin or self glucosylating protein (SGP) is a starch primer.
 CC It can be purified from the endosperm of a US inbred line (B73) of
 CC Zea mays. The protein was then digested with trypsin in an
 CC enzyme/substrate ratio of 1:25 for 15 hr. The tryptic digest was
 CC fractionated. Nine of the peptide peaks (labelled P1-P9) were
 CC further purified. Upon refractionated of peaks P1-P9 a number of
 CC tryptic peptide peaks were obt'd. Nine of the purified tryptic
 CC peptides were chosen for AA sequence analyses (labelled T1-T9).
 CC The sequences of these peptides is given in AAR49882-90. The
 CC radiolabelled peptide T2 is a proteolytic fragment of the larger
 CC radiolabelled peptide T6.
 XX
 XX
 SQ Sequence 28 AA;
 Query Match 100.0%; Score 23; DB 15; Length 28;
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DYDA 4
 Db ||||
 17 DYDA 20
 RESULT 8
 ABB42618
 ID ABB42618 standard; Peptide; 38 AA.
 XX
 AC ABB42618;
 XX
 DT 04-FEB-2002 (first entry)
 XX
 DE Peptide #10124 encoded by human foetal liver single exon probe.
 DE
 XX Human; foetal liver; gene expression; single exon nucleic acid probe.
 KW Homo sapiens.
 OS
 PN WO200157277-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US00669.
 PF
 XX 04-FEB-2000; 2000US-0180312.
 PR
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 XX Penn SG, Hanzel DK, Chen W, Rank DR;
 PI WPI; 2001-483447/52.
 XX
 DR Human genome-derived single exon nucleic acid probes useful for
 DR analyzing gene expression in human fetal liver -
 XX
 PT Claim 27; SEQ ID NO 35253; 639pp + sequence listing; English.
 PS
 XX The invention relates to a single exon nucleic acid probe for
 CC measuring human gene expression in a sample derived from human foetal
 CC liver. The single exon nucleic acid probes may be used for predicting,
 CC measuring and displaying gene expression in samples derived from human
 CC foetal liver. The present sequence is a peptide encoded by a single exon
 CC nucleic acid probe of the invention.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly

CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 38 AA;

Query Match 100.0%; Score 23; DB 22; Length 38;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DYDA 4
||||
Db 19 DYDA 22

RESULT 9

AAM63509
ID AAM63509 standard; Protein; 38 AA.

XX AAM63509;

XX 05-NOV-2001 (first entry)

XX Human brain expressed single exon probe encoded protein SEQ ID NO: 35614.

XX Human; brain expressed exon; gene expression analysis; probe;
KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
KW epilepsy; cancer.

XX Homo sapiens.

XX WO200157275-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US00667.

XX 04-FEB-2000; 2000US-0180312.

XX 26-MAY-2000; 2000US-0207456.

XX 30-JUN-2000; 2000US-0608408.

XX 03-AUG-2000; 2000US-0632366.

XX 21-SEP-2000; 2000US-0234687.

XX 27-SEP-2000; 2000US-0236359.

XX 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-483446/52.

XX Single exon nucleic acid probes for analyzing gene expression in human

XX brains -
XX Example 4; SEQ ID NO: 35614; 650pp + Sequence Listing; English.

XX The present invention provides a number of single exon nucleic acid

XX probes which are derived from genomic sequences expressed in the human
XX brain. They can be used to measure gene expression in brain cell samples,
XX which may enable the diagnosis and improved treatment of nervous system
XX diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
XX epilepsy and cancers. The present sequence is a protein encoded by one of
XX the probes of the invention.

XX Sequence 38 AA;

Qy 1 DYDA 4

Db 19 DYDA 22

RESULT 10

AAM76323
ID AAM76323 standard; Protein; 38 AA.

XX AAM76323;

XX 06-NOV-2001 (first entry)

XX Human bone marrow expressed probe encoded protein SEQ ID NO: 36629.

XX Human; bone marrow expressed exon; gene expression analysis; probe;
KW microarray; cancer; leukaemia; lymphoma; myeloma.

XX Homo sapiens.

XX WO200157276-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US00668.

XX 04-FEB-2000; 2000US-0180312.

XX 26-MAY-2000; 2000US-0207456.

XX 30-JUN-2000; 2000US-0608408.

XX 03-AUG-2000; 2000US-0632366.

XX 21-SEP-2000; 2000US-0234687.

XX 27-SEP-2000; 2000US-0236359.

XX 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-488900/53.

XX Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human bone marrow -

XX Example 4; SEQ ID NO: 36629; 658pp + Sequence Listing; English.

XX The present invention provides a number of single exon nucleic acid

XX probes which are derived from genomic sequences expressed in the human

XX bone marrow. They can be used to measure gene expression in bone marrow

XX samples, which may enable the improved diagnosis and treatment of cancers

XX such as lymphoma, leukaemia and myeloma. The present sequence is a

XX protein encoded by one of the probes of the invention.

XX Sequence 38 AA;

Qy 1 DYDA 4

Db 19 DYDA 22

RESULT 11

AAM36432
ID AAM36432 standard; Protein; 38 AA.

XX AAM36432;

XX 17-OCT-2001 (first entry)

XX Peptide #10469 encoded by probe for measuring placental gene expression.

XX Probe; microarray; human; placenta; antenatal diagnosis;

XX genetic disorder.

XX Homo sapiens.

```
PN WO200157272-A2.
XX 09-AUG-2001.
XX 30-JAN-2001; 2001WO-US00663.
XX 04-FEB-2000; 2000US-0180312.
XX 26-MAY-2000; 2000US-0207456.
XX 30-JUN-2000; 2000US-0608408.
XX 03-AUG-2000; 2000US-0632366.
XX 21-SEP-2000; 2000US-0234687.
XX 27-SEP-2000; 2000US-0236359.
XX 04-OCT-2000; 2000GB-0024263.
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488897/53.
XX Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human placenta -
XX Claim 27; SEQ ID NO 36701; 654pp; English.
XX
XX The present invention relates to single exon nucleic acid probes (SENP:
XX see AAI31315-AAI57546). The present sequence is a peptide encoded by one
XX such probe. The probes are useful for producing a microarray for
XX predicting, measuring and displaying gene expression in samples derived
XX from human placenta. The probes are useful for antenatal diagnosis of
XX human genetic disorders.
XX
XX SQ Sequence 38 AA;
Query Match 100.0%; Score 23; DB 22; Length 38;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 DYDA 4
Db 19 DYDA 22
||||
RESULT 12
AAG10870
ID AAG10870 standard; Protein; 41 AA.
XX
XX AAG10870;
XX
XX 17-OCT-2000 (first entry)
XX
XX Arabidopsis thaliana protein fragment SEQ ID NO: 9365.
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX
XX Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
XX 09-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0125788.
XX 25-MAR-1999; 99US-0126264.
XX 29-MAR-1999; 99US-0126785.
XX 01-APR-1999; 99US-0127462.
XX 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 23-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
```

```
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145119.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148563.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150360.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151063.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158234.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
Query Match 100.0%; Score 23; DB 21; Length 41;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DYDA 4
Db 28 DYDA 31
RESULT 13
AAG54078
ID AAG54078 standard; Protein; 41 AA.
XX
AC AAG54078;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 68915.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 16-APR-1999; 99US-0128714.
PR 18-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
```

PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139452.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 30-JUN-1999; 99US-0140991.
PR 01-JUL-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142820.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.

PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155569.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.

PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 100.0%; Score 23; DB 21; Length 41;
 Best Local Similarity 100.0%; Pred. No. 2e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
 ||||
 Db 28 DYDA 31

RESULT 14
 AAW79339
 ID AAW79339 standard; Protein; 44 AA.
 AC AAW79339;
 XX
 DT 24-NOV-1998 (first entry)
 DE Staphylococcus aureus haein lipioic acid synthetase.
 KW Haemophilus influenzae; haein lipioic acid synthetase;
 KW lipocate biosynthesis protein A; lipA; treatment; prevention;
 KW bacterial infection; Helicobacter pylori; vaccine.
 XX
 OS Staphylococcus aureus.
 XX
 PN W09823738-A2.
 XX
 PD 04-JUN-1998.
 XX
 PF 24-NOV-1997; 97WO-US22092.
 XX
 PR 25-NOV-1996; 96US-0031469.
 XX
 PA (SMIK) SMITHKLINE BEECHAM CORP.
 XX
 PI Warren RL;
 XX
 DR WPI; 1998-322718/28.
 DR N-PSDB; AAV59884.
 XX
 PT New nucleic acid from Staphylococcus aureus NCIMB 40771 - useful
 for, e.g. diagnosis, prevention and treatment of bacterial
 infection(s)
 PT
 XX
 PS Claim 5; Pages 51; 114pp; English.
 XX
 CC AAW71314 and AAW79337-39 represent Staphylococcus aureus WCHU (NCIMB
 40771) proteins that have homology to Haemophilus influenzae haein
 lipioic acid synthetase (lipocate biosynthesis protein A (lipA) homologue.
 CC The protein is used to generate antibodies and to screen for
 CC antimicrobials. The products are used to treat or prevent bacterial
 CC infections, particularly where caused by S. aureus but also
 CC against Helicobacter pylori. Particular applications are to treat

CC subjects before surgery or insertion of an in-dwelling device
 CC (alternatively the device itself is impregnated before placement). The
 CC nucleic acid sequence is used as sources of antisense sequences (for
 CC therapeutic use) or regulatory elements for controlling expression of
 CC bacterial genes, and for antibacterial screening. The protein can be
 CC also used as a vaccine.
 XX
 SQ Sequence 44 AA;
 Query Match 100.0%; Score 23; DB 19; Length 44;
 Best Local Similarity 100.0%; Pred. No. 2.1e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
 ||||
 Db 9 DYDA 12

RESULT 15
 AAW62275
 ID AAW62275 standard; peptide; 45 AA.
 XX
 AC AAW62275;
 XX
 DT 22-SEP-1998 (first entry)
 DE Streptococcus pneumoniae pspA protein N-terminal peptide.
 XX
 KW Streptococcus pneumoniae strain Rxl; pspA; immunoprotective; immunogen;
 KW pneumococcal surface protein A; cholera toxin B subunit; fusion protein;
 KW antigenic.
 XX
 OS Streptococcus pneumoniae.
 XX
 PN US5753463-A.
 XX
 PD 19-MAY-1998.
 XX
 PF 06-JUN-1995; 95US-0469434.
 XX
 PR 12-FEB-1992; 92US-0835698.
 PR 15-FEB-1991; 91US-0656773.
 PR 03-JUN-1993; 93US-0072065.
 PR 06-JUN-1995; 95US-0469434.
 XX
 PA (UABR-) UAB RES FOUND.
 XX
 PI Briles DE, Yother JL;
 XX
 DR WPI; 1998-311399/27.
 XX
 PT Truncated pneumococcal surface protein and cholera toxin B sub-unit
 PT fusion protein - useful as an immunogen against Streptococcus
 PT pneumoniae
 XX
 PS Example 5; Fig 2; 22pp; English.
 XX
 CC The present sequence represents an N-terminal peptide of pneumococcal
 CC surface protein A (pspA) protein from Streptococcus pneumoniae. A
 CC recombinant DNA molecule has been developed which encodes a fusion
 CC protein comprising a truncated form of pspA and cholera toxin B subunit
 CC (CTB), where the DNA molecule comprises a nucleotide sequence encoding
 CC the truncated pspA linked by an in-frame genetic fusion to a ctxB gene,
 CC and where the truncated pspA contains immunoprotective epitopes and up
 CC to 90% of the whole pspA protein, except for the cell membrane anchor
 CC region. The fusion protein is useful for providing an immunogen to
 CC protect neonates and children against S.pneumoniae. Most antigenic
 CC proteins of this strain are not immunogenic enough to provide protection.
 CC The antigenic epitopes of the fusion protein are directed against
 CC capsular polysaccharide antigens of S.pneumoniae, specifically it
 CC contains the protective epitopes of pspA. The protein can also be used
 CC in solid-phase immunoadsorbent assays, since it is readily bound to
 CC supports coated with monosialoganglioside GM1. The fusion protein is

CC more immunogenic against S.pneumoniae than using PspA alone as the
CC immunogen.

SQ Sequence 45 AA;

Query Match 100.0%; Score 23; DB 19; Length 45;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
|||||
Db 14 DYDA 17

RESULT 16

AAAY41839
ID AAY41839 standard; Protein; 45 AA.

XX AAY41839;

AC AAY41839;

DT 08-DEC-1999 (first entry)

DE Streptococcus pneumoniae Rxl PspA N-terminal sequence #1.

XX Streptococcus pneumoniae Rxl; PspA; immunoprotective; vaccine;

KW diagnosis; infection; pneumococcal surface protein A.

XX Streptococcus pneumoniae.

OS
XX US5965400-A.

PN 12-OCT-1999.

PD 23-MAY-1994; 94US-0247491.

PF 12-FEB-1992; 92US-0835698.

PR 15-FEB-1991; 91US-0656773.

PS (UABR-) UAB RES FOUND.

XX Yother JL, Briles DE;

DR WPI; 1999-579913/49.

XX DNA encoding a truncated pneumococcal surface protein A used in the
PT development of pneumococcal infections -

PS Disclosure; Fig 2; 27pp; English.

XX The present sequence represents a Streptococcus pneumoniae Rxl
CC immunoprotective Pneumococcal surface protein A (PspA) N-terminal
CC sequence. The present invention also describes a method of forming
CC the immunoprotective truncated PspA, comprising incorporating a vector
CC comprising the isolated DNA molecule encoding PspA (I), into a
CC bacterium via transformation. (I) is used to design primers which are
CC capable of detecting a large number of S. pneumoniae strains, which in
CC turn can be used to diagnose pneumococcal infection in mammals (e.g.
CC humans), independent of the strain which has caused it. The PspA protein
CC is used to develop a vaccine against pneumococcal infection comprising,
CC as an immunologically-active component, a live attenuated or killed
CC bacteria containing a gene coding for the truncated form of PspA.

SQ Sequence 45 AA;

Query Match 100.0%; Score 23; DB 20; Length 45;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
|||||
Db 14 DYDA 17

RESULT 17

AAW92457
ID AAW92457 standard; Protein; 45 AA.

XX AAW92457;

AC AAW92457;

DT 21-APR-1999 (first entry)

DE S. pneumoniae truncated pspA protein fragment #1.

XX PspA; pneumococcal surface protein A; immunoprotection; detection; GMI;
KW solid phase immunosorbent assay; epitope; cell membrane anchor region;
KW cholera toxin B subunit; CTB; monosialoganglioside; fusion protein.

XX Streptococcus pneumoniae.

OS US5871943-A.

PN 16-FEB-1999.

PD 06-JUN-1995; 95US-0468718.

PF 12-FEB-1992; 92US-0835698.

PR 15-FEB-1991; 91US-0656773.

PR 03-JUN-1993; 93US-0072068.

PR 06-JUN-1995; 95US-0468718.

XX (UABR-) UAB RES FOUND.

PA Briles DE, Yother JL;

XX WPI; 1999-166635/14.

DR Immunosorbent assay for pneumococcal surface protein A antigen or
XX antibody - for diagnosis of infection by Streptococcus pneumoniae

PS Disclosure; Fig 2; 24pp; English.

XX This sequence represents a fragment of the truncated form of the
CC Streptococcus pneumoniae PspA protein which is used in a solid phase
CC immunosorbent assay for detecting a PspA (pneumococcal surface protein A)
CC antibody and antigen. This truncated protein contains the
CC immunoprotective epitopes of the complete protein (up to 90% of PspA but
CC excludes the cell-membrane anchor region) fused to the B subunit of
CC cholera toxin (CTB) which is bound to monosialoganglioside (GMI) coated
CC on the substrate. The use of a fusion between truncated PspA and cholera
CC toxin B subunit (CTB) allows the support to be coated without having to
CC isolate PspA fragments since CTB binds specifically to the GMI coating
CC the solid support.

SQ Sequence 45 AA;

Query Match 100.0%; Score 23; DB 20; Length 45;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
|||||
Db 14 DYDA 17

RESULT 18

AAW04953
ID AAW04953 standard; peptide; 49 AA.

XX AAW04953;

AC AAW04953;

DT 21-MAY-1997 (first entry)

DE Partial D. gossypina endoglucanase.

XX Cellulolytic enzyme; endoglucanase; hydrolysis; cellulose; microorganism;
KW plant cellulase; catalytic region; textile; backstaining; bio-polishing;

stone-washing; cellulosic fabric; colour clarification; defibrillation;
cell wall degradation; paper pulp; debarking; fibre modification;
enzymatic de-inking; drainage improvement.

Diplodia gossypina.

WO9629397-A1.

26-SEP-1996.

18-MAR-1996; 96WO-DK00105.

12-FEB-1996; 96DK-0000137.

17-MAR-1995; 95DK-0000272.

08-AUG-1995; 95DK-0000885.

08-AUG-1995; 95DK-0000886.

08-AUG-1995; 95DK-0000887.

08-AUG-1995; 95DK-0000888.

(NOVO) NOVO-NORDISK AS.

Andersen LN, Ihara M, Kauppinen MS, Lange L, Lassen SF;
Nielsen RI, Schuelein M, Takagi S;

WPI; 1996-443173/44.

N-PSDB; AAT39091.

New endo-glucanase enzyme preparations - contg. conserved catalytic
regions, useful for treating fabrics, textiles, plant material or
paper pulp

Example 5; Page 217; 316pp; English.

AAW04938-W04963 represent partial endoglucanase sequences. These
peptides are examples of fragments of the enzymes of the invention (see
AAW04925 for an example of a full length enzyme of the invention). The
full length enzymes possess cellulolytic (particularly endoglucanase)
activity. Cellulolytic enzymes are involved in the hydrolysis of cellulose,
and are synthesised by a large number of microorganisms and plants. The
enzymes of the invention containing the conserved catalytic regions (such
as AAW04913) exhibit improved performance, e.g. 50 times higher
performance, compared to multiple domain enzymes. The enzymes can be
used for the treatment of fabrics or textiles, preferably for preventing
backstaining, for bio-polishing or for stone-washing cellulosic fabric.
They can also be used to provide colour clarification for laundry. The
enzymes can also be used for the degradation or modification of plant
material, such as cell walls. They can also be used in the treatment of
paper pulp preferably for debarking, defibrillation, fibre modification,
enzymatic de-inking or drainage improvement.

Sequence 49 AA;

Query Match 100.0%; Score 23; DB 17; Length 49;
Best Local Similarity 100.0%; Pred. NO. 2.4e-02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
|||||
Db 11 DYDA 14

RESULT 19
ABBI1106
ID ABBI1106 standard; peptide; 58 AA.
XX
AC ABBI1106;
XX
DT 11-JAN-2002 (first entry)
XX
DE Human apoptosis regulatory kinase homologue, SEQ ID NO:1476.
XX
KW Human; cytokine; cell proliferation; cell differentiation; growth factor;
haematopoiesis regulation; tissue growth; immunomodulator; activin;

inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;
proliferation; metastasis; cancer; tumour; haematopoietic disorder;
myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
chronic inflammatory condition; proliferative retinopathy;
atherosclerosis; coronary heart disease; arterial ischaemia;
bone disorder; osteoporosis; vascular growth disorder;
tissue regeneration; wound healing; infection; immune disorder;
cell culture; drug screening; gene therapy; antiinflammatory;
antiasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;
cytostatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;
antifungal; vulnery; antiulcer.

Homo sapiens.

WO200157188-A2.

09-AUG-2001.

05-FEB-2001; 2001WO-US03800.

03-FEB-2000; 2000US-0496914.

27-APR-2000; 2000US-0560875.

(HYSE-) HYSEQ INC.

Tang YT, Liu C, Drmanac RT;

WPI; 2001-457740/49.

N-PSDB; ABA08350.

Human proteins and DNA encoding sequences useful for preventing,
treating or ameliorating a medical condition in a mammalian subject
e.g. arthritis and cancer -

Claim 20; Page 146; 1963pp; English.

Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and
sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
invention also relates to vectors and recombinant host cells comprising a
nucleotide of the invention, methods of producing the novel polypeptides,
antibodies against the polypeptides, methods of detecting the nucleotides
or polypeptides in a sample, and methods of identifying compounds which
bind to polypeptides of the invention. Although novel, many of the
polypeptides of the invention have homology to known proteins, thereby
giving an insight into their probable biological activities, and hence
potential therapeutic applications. The polypeptides of the invention may
have various activities, including cytokine, cell proliferation or cell
differentiation activities; stem cell growth factor activity;
haematopoiesis regulatory activity; tissue growth activity;
immunomodulatory activity; activin- or inhibin-related activities;
chemotactic or chemokinetic activities; haemostatic, thrombotic or
thrombolytic activities; receptor or ligand activities; or may be
involved in oncogenesis, cancer cell proliferation or metastasis.
Depending on their biological activities, polypeptides and nucleotides of
the invention are useful for preventing, treating or ameliorating medical
conditions, e.g., by protein or gene therapy. Such conditions include
cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell
disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
proliferative retinopathy, atherosclerosis, coronary heart disease,
arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
vascular growth. Polypeptides involved with tissue regeneration and
repair (or nucleic acids encoding them) may be used to promote wound
healing (e.g., of burns, incisions and ulcers), while those with
immunomodulatory activities may be used in the treatment of viral,
bacterial and fungal infections in addition to immune disorders.
Polypeptides with growth factor activity may be used in cell cultures to
promote cell growth. For example, such polypeptides may be used to
manipulate stem cells in culture to give rise to neuroepithelial cells
that can be used to augment or replace cells damaged by illness,
autoimmune disease or accidental damage. The polypeptides and nucleotides
may also be used in the diagnosis of the above conditions, and in drug
screening techniques. The present sequence represents a novel human
polypeptide of the invention.

```
XX SQ Sequence 58 AA;
Query Match 100.0%; Score 23; DB 22; Length 58;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
Db 20 DYDA 23

RESULT 20
AAU67058
ID AAU67058 standard; Protein; 61 AA.
XX
AC AAU67058;
XX
DT 27-FEB-2002 (first entry)
XX
DE Propionibacterium acnes immunogenic protein #27954.
XX
KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant.
XX
OS Propionibacterium acnes.
XX
PN WO200181581-A2.
XX
PD 01-NOV-2001.
XX
PF 20-APR-2001; 2001WO-US12865.
XX
PR 21-APR-2000; 2000US-199047P.
PR 02-JUN-2000; 2000US-208841P.
PR 07-JUL-2000; 2000US-216747P.
XX
PA (CORI-) CORIYA CORP.
XX
PI Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX
DR WPI; 2001-616774/71.
DR N-PSDB; AAS59776.
XX
PT Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris -
XX
PS Example 1; SEQ ID No 28253; 1069pp; English.
XX
CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA).
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
```

```
XX SQ Sequence 61 AA;
Query Match 100.0%; Score 23; DB 22; Length 61;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
Db 46 DYDA 49

RESULT 21
AAW62591
ID AAW62591 standard; peptide; 64 AA.
XX
AC AAW62591;
XX
DT 17-SEP-1998 (first entry)
XX
DE Peptide having affinity for a malarial Duffy-binding ligand.
XX
KW Duffy protein; DP; binding affinity; malarial Duffy-binding ligand;
KW MDBL; human; gp-Fy-alpha; gp-Fy-beta; malaria; inhibit; invasion;
KW red blood cell; malarial parasite; merozoite; Plasmodium vivax;
KW P. knowlesi; treatment; regulation; erythrocyte; neural; renal function.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9821235-A1.
XX
PD 22-MAY-1998.
XX
PF 14-NOV-1997; 97WO-US21063.
XX
PR 15-NOV-1996; 96US-0749526.
XX
PA (NYBL-) NEW YORK BLOOD CENT INC.
XX
PI Chaudhuri A, Pogo OA;
XX WPI; 1998-297863/26.
XX
PT New peptide(s) from Duffy protein that bind to malarial Duffy
PT binding ligand - used to protect against infection by malarial
PT parasites, particularly Plasmodium vivax or Plasmodium knowlesi
XX
PS Claim 3; Page 4; 39pp; English.
XX
CC Peptides AAW62589-91 have amino acid sequences characteristic of a Duffy
CC protein (DP) and specific binding affinity for malarial Duffy-binding
CC ligand (MDBL). The human Duffy protein can be a gp-Fy-alpha or a
CC gp-Fy-beta protein. The peptides are used to protect against malaria.
CC They specifically inhibit the invasion of red blood cells by malarial
CC parasites, specifically merozoites of Plasmodium vivax and P. knowlesi.
CC The peptides can also be used to generate antibodies, complementary
CC peptides and drugs modelled on their tertiary structures, and are all
CC potentially useful in treatment of malaria and in the regulation of
CC essential erythrocyte, neural and renal functions.
XX
SQ Sequence 64 AA;
Query Match 100.0%; Score 23; DB 19; Length 64;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
Db 40 DYDA 43

RESULT 22
```

AAW73956
ID AAR73956 standard; Peptide; 66 AA.

XX AC AAR73956;

XX DT 22-NOV-1995 (first entry)

XX DE Duffy blood group gpD protein antigen N-terminal peptide.

XX KW Duffy blood group; malaria; Plasmodium vivax; glycoprotein; gpD;

XX KW cytokine receptor.

XX OS Homo sapiens.

XX PN WO9511040-A.

XX PD 27-APR-1995.

XX PF 20-OCT-1994; 94WO-US12028.

XX PR 21-OCT-1993; 93US-0140797.

XX PA (NYBL-) NEW YORK BLOOD CENT INC.

XX PI Chaudhuri A, Pogo AO, Pogo A;

XX DR WPI; 1995-193792/25.

XX PT DNA encoding gpD protein component to Duffy blood group antigen - also
PT proteins, peptide(s) and vaccines for the therapy of malaria, and the
PT regulation of erythrocyte, neural and renal functions

XX PS Claim 5; Page 10; 49pp; English.

XX CC The N-terminal region of the gpD protein has been identified as
CC being involved in the interaction of the malaria parasite with the
CC red cell. This peptide binds the Rubinstein antibody which is a
CC monoclonal antibody specific for the gpD protein. The fact that
CC this peptide along with those of AAR73957-58 bind the Rubinstein
CC antibody whereas AAR73959-61 do not suggest that the junction
CC between the C-terminal of peptide AAR73961 and the N-terminal
CC of AAR73959 is important for binding. This peptide binds the
CC parasite in vivo and, therefore, can be used as the immunogen in
CC a vaccine against malaria.

XX SQ Sequence 66 AA;

Query Match 100.0%; Score 23; DB 16; Length 66;

Best Local Similarity 100.0%; Pred. No. 3.3e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4

DB 42 DYDA 45

RESULT 23

AAW62590

ID AAW62590 standard; peptide; 66 AA.

XX AC AAW62590;

XX DT 17-SEP-1998 (first entry)

XX DE Peptide having affinity for a malarial Duffy-binding ligand.

XX KW Duffy protein; DP; binding affinity; malarial Duffy-binding ligand;

XX KW MBDL; human; gp-Fy-alpha; gp-Fy-beta; malaria; inhibit; invasion;

XX KW red blood cell; malarial parasite; merozoite; Plasmodium vivax;

XX KW P. knowlesi; treatment; regulation; erythrocyte; neural; renal function.

XX OS Synthetic.

XX OS Homo sapiens.

XX WO9821235-A1.

XX PD 22-MAY-1998.

XX PF 14-NOV-1997; 97WO-US21063.

XX PR 15-NOV-1996; 96US-0749526.

XX PA (NYBL-) NEW YORK BLOOD CENT INC.

XX PI Chaudhuri A, Pogo OA;

XX DR WPI; 1998-297863/26.

XX PT New peptide(s) from Duffy protein that bind to malarial Duffy
PT binding ligand - used to protect against infection by malarial
PT parasites, particularly Plasmodium vivax or Plasmodium knowlesi

XX PS Claim 3; Page 4; 39pp; English.

XX CC Peptides AAW62589-91 have amino acid sequences characteristic of a Duffy
CC protein (DP) and specific binding affinity for malarial Duffy-binding
CC ligand (MBDL). The human Duffy protein can be a gp-Fy-alpha or a
CC gp-Fy-beta protein. The peptides are used to protect against malaria.
CC They specifically inhibit the invasion of red blood cells by malarial
CC parasites, specifically merozoites of Plasmodium vivax and P. knowlesi.
CC The peptides can also be used to generate antibodies, complementary
CC peptides and drugs modelled on their tertiary structures, and are all
CC potentially useful in treatment of malaria and in the regulation of
CC essential erythrocyte, neural and renal functions.

XX SQ Sequence 66 AA;

Query Match 100.0%; Score 23; DB 19; Length 66;

Best Local Similarity 100.0%; Pred. No. 3.3e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4

DB 42 DYDA 45

RESULT 24

AAW54861

ID AAW54861 standard; Protein; 66 AA.

XX AC AAW54861;

XX DT 11-SEP-1998 (first entry)

XX DE Homo sapiens anti-Fy6 antibody specifically binding peptide.

XX KW gp-Fy protein; Fyb71-81; duffy blood group; antigen; alpha; beta;
XX KW alternative splicing; RBC; red blood cell; malaria; treatment.

XX OS Homo sapiens.

XX PN WO9821224-A1.

XX PD 22-MAY-1998.

XX PF 14-NOV-1997; 97WO-US21067.

XX PR 15-NOV-1996; 96US-0749543.

XX PA (NYBL-) NEW YORK BLOOD CENT INC.

XX PI Chaudhuri A, Pogo OA;

XX DR WPI; 1998-297854/26.

XX PT Nucleic acid encoding gp-Fy, Duffy antigen proteins - used to

PT prevent vivax malaria and to regulate erythrocyte, neural or renal
function

PS Claim 10; Page 13; 87pp; English.

XX The sequence is that of a peptide which specifically binds
CC the anti-Fy6 antibody, which blocks penetration of P.vivax
CC merozoites into human erythrocytes in vitro. The peptide may
CC be useful in preventing malaria and may also be
CC used to generate antibodies, complementary peptides and drugs
CC modelled on its tertiary structure.

XX Sequence 66 AA;

Query Match 100.0%; Score 23; DB 19; Length 66;

Best Local Similarity 100.0%; Pred. No. 3.3e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4

||||

Db 42 DYDA 45

RESULT 25

AAB12795

ID AAB12795 standard; protein; 67 AA.

XX

AC AAB12795;

DT 23-NOV-2000 (first entry)

XX Clostridium thermocellum amino acid sequence SEQ ID NO:8.

XX Clostridium; multi-enzyme complex; immobilised body; carrier; food;

KW cohensin-dockerin; cellulose bound protein; medicine.

XX Clostridium thermocellum.

OS JP2000157282-A.

XX 13-JUN-2000.

PF 30-NOV-1998; 98JP-0338671.

XX 30-NOV-1998; 98JP-0338671.

XX (TOYU) TOYOTA CHUO KENKYUSHO KK.

XX WPI; 2000-468205/41.

XX Immobilized multi-enzyme complex, optionally modified, and a method
PT for its preparation

XX Disclosure; Page 10; 11pp; Japanese.

XX The present invention describes an optionally immobilised multi-enzyme
CC complex prepared by specific binding with cohensin-dockerin on cellulose
CC bound protein. An optionally immobilised protein carrier molecule has
CC plural repetitive enzyme bound domains (C) capable of binding with a
CC carrier molecular bound domain (D) of cellulose degrading enzyme, in
CC which the cellulose bound protein contains intact cellulose bound domain
CC (CBD) or the carrier having replaced plural synaptonemal enzyme domains
CC (e.g. C1, C2) with different binding properties. The immobilised
CC multi-enzyme complex is useful for various medicine and food industries.
CC AAB12788 to AAB12796 represent Clostridium protein sequences given in
CC the sequence listing from the present invention, but not defined further
CC within the specification.

XX Sequence 67 AA;

Query Match

Best Local Similarity 100.0%; Score 23; DB 21; Length 67;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4

||||

Db 63 DYDA 66

RESULT 26

ABG23865

ID ABG23865 standard; Protein; 68 AA.

XX ABG23865;

XX 18-FEB-2002 (first entry)

XX Novel human diagnostic protein #23856.

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;

KW food supplement; medical imaging; diagnostic; genetic disorder.

XX Homo sapiens.

XX WO200175067-A2.

XX 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US08631.

XX 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

XX N-PSDB; AAS88052.

XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity

XX Claim 20; SEQ ID No 54224; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG0010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 68 AA;

Query Match 100.0%; Score 23; DB 22; Length 68;

Best Local Similarity 100.0%; Pred. No. 3.4e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4

Db 55 DYDA 58
||||

RESULT 27

ABB42747
ID ABB42747 standard; Peptide; 69 AA.

XX AC ABB42747;
XX DT

XX 04-FEB-2002 (first entry)

XX Peptide #10253 encoded by human foetal liver single exon probe.

XX Human; foetal liver; gene expression; single exon nucleic acid probe.

XX Homo sapiens.
XX PN WO200157277-A2.
XX PD 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US00669.

XX 04-FEB-2000; 2000US-0180312.
XX PR 26-MAY-2000; 2000US-0207456.
XX PR 30-JUN-2000; 2000US-0608408.
XX PR 03-AUG-2000; 2000US-0632366.
XX PR 21-SEP-2000; 2000US-0234687.
XX PR 27-SEP-2000; 2000US-0236359.
XX PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-483447/52.

XX Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human fetal liver -

XX Claim 27; SEQ ID NO 35382; 639pp + sequence listing; English.

XX The invention relates to a single exon nucleic acid probe for
XX measuring human gene expression in a sample derived from human foetal
XX liver. The single exon nucleic acid probes may be used for predicting,
XX measuring and displaying gene expression in samples derived from human
XX foetal liver. The present sequence is a peptide encoded by a single exon
XX nucleic acid probe of the invention.
XX Note: The sequence data for this patent did not form part of the
XX printed specification, but was obtained in electronic format directly
XX from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 69 AA;

Query Match 100.0%; Score 23; DB 22; Length 69;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 DYDA 4
||||

Db 34 DYDA 37

RESULT 28

ABB26046
ID ABB26046 standard; Protein; 69 AA.

XX AC ABB26046;
XX DT

XX 23-JAN-2002 (first entry)

XX Protein #8045 encoded by probe for measuring heart cell gene expression.

XX Human; gene expression; heart; microarray; vascular system;
KW cardiovascular disease; hypertension; cardiac arrhythmia;
KW congenital heart disease.

XX Homo sapiens.
XX PN WO200157274-A2.
XX PD 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US00666.

XX 04-FEB-2000; 2000US-0180312.
XX PR 26-MAY-2000; 2000US-0207456.
XX PR 30-JUN-2000; 2000US-0608408.
XX PR 03-AUG-2000; 2000US-0632366.
XX PR 21-SEP-2000; 2000US-0234687.
XX PR 27-SEP-2000; 2000US-0236359.
XX PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488899/53.

XX Single exon nucleic acid probes for analyzing gene expression in human
XX hearts -

XX Claim 15; SEQ ID NO 27816; 530pp; English.

XX The present invention relates to single exon nucleic acid probes for
XX measuring human gene expression in a sample derived from human heart (see
XX ABA21535-ABA41305). The present sequence is a protein encoded by one such
XX probe. The probes may be used for predicting, measuring and displaying
XX gene expression in samples derived from the human heart via microarrays.
XX By measuring gene expression, the probes are useful for predicting,
XX diagnosing, grading, staging, monitoring and prognosing diseases of the
XX human heart and vascular system e.g. cardiovascular disease,
XX hypertension, cardiac arrhythmias and congenital heart disease.
XX Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 69 AA;

Query Match 100.0%; Score 23; DB 22; Length 69;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 DYDA 4
||||

Db 34 DYDA 37

RESULT 29

AAM63638
ID AAM63638 standard; Protein; 69 AA.

XX AC AAM63638;
XX DT

XX 05-NOV-2001 (first entry)

XX Human brain expressed single exon probe encoded protein SEQ ID NO: 35743.

XX Human; brain expressed exon; gene expression analysis; probe;
KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
KW epilepsy; cancer.

XX Homo sapiens.
XX PN WO200157275-A2.

XX 09-AUG-2001.
PD
XX
PF 30-JAN-2001; 2001WO-US00667.
XX
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-483446/52.
DR
XX
XX Single exon nucleic acid probes for analyzing gene expression in human brains -
PT
PT
XX
PS Example 4; SEQ ID NO: 35743; 650pp + Sequence Listing; English.
XX
XX The present invention provides a number of single exon nucleic acid probes which are derived from genomic sequences expressed in the human brain. They can be used to measure gene expression in brain cell samples, which may enable the diagnosis and improved treatment of nervous system diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia, epilepsy and cancers. The present sequence is a protein encoded by one of the probes of the invention.
XX
XX
SQ Sequence 69 AA;
Query Match 100.0%; Score 23; DB 22; Length 69;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 DYDA 4
Db 34 DYDA 37
|||||
RESULT 30
AAM76452
ID AAM76452 standard; Protein; 69 AA.
AC
XX
AC AAM76452;
XX
DT 06-NOV-2001 (first entry)
XX
XX Human bone marrow expressed probe encoded protein SEQ ID NO: 36758.
DE
XX Human; bone marrow expressed exon; gene expression analysis; probe; microarray; cancer; leukaemia; lymphoma; myeloma.
KW
KW Homo sapiens.
XX
OS
XX WO200157276-A2.
PN
XX
PD 09-AUG-2001.
XX
XX
PF 30-JAN-2001; 2001WO-US00668.
XX
XX 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;
PI
XX WPI; 2001-488900/53.
DR
XX Human genome-derived single exon nucleic acid probes useful for analyzing gene expression in human bone marrow -
PT
PT
XX
XX Example 4; SEQ ID NO: 36758; 658pp + Sequence Listing; English.
PS
XX The present invention provides a number of single exon nucleic acid probes which are derived from genomic sequences expressed in the human bone marrow. They can be used to measure gene expression in bone marrow samples, which may enable the improved diagnosis and treatment of cancers such as lymphoma, leukaemia and myeloma. The present sequence is a protein encoded by one of the probes of the invention.
XX
XX
SQ Sequence 69 AA;
Query Match 100.0%; Score 23; DB 22; Length 69;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 DYDA 4
Db 34 DYDA 37
|||||
RESULT 31
AAM36561
ID AAM36561 standard; Protein; 69 AA.
AC
XX
AC AAM36561;
XX
DT 17-OCT-2001 (first entry)
XX
XX Peptide #10598 encoded by probe for measuring placental gene expression.
DE
XX Probe; microarray; human; placenta; antenatal diagnosis;
KW genetic disorder.
KW
XX Homo sapiens.
OS
XX
XX WO200157272-A2.
PN
XX
PD 09-AUG-2001.
XX
XX
PF 30-JAN-2001; 2001WO-US00663.
XX
XX 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
PA
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
PI
XX WPI; 2001-488897/53.
DR
XX Human genome-derived single exon nucleic acid probes useful for analyzing gene expression in human placenta -
PT
PT
XX
XX Claim 27; SEQ ID No 36830; 654pp; English.
PS
XX The present invention relates to single exon nucleic acid probes (SENP: see AAI31315-AAI57546). The present sequence is a peptide encoded by one such probe. The probes are useful for producing a microarray for predicting, measuring and displaying gene expression in samples derived from human placenta. The probes are useful for antenatal diagnosis of

CC human genetic disorders.

XX Sequence 69 AA;

Query Match 100.0%; Score 23; DB 22; Length 69;

Best Local Similarity 100.0%; Pred. No. 3.5e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4

||||

Db 34 DYDA 37

RESULT 32

ABG45716

ID ABG45716 standard; Peptide; 69 AA.

AC ABG45716;

XX 19-AUG-2002 (first entry)

XX Human peptide encoded by genome-derived single exon probe SEQ ID 35381.

XX Human; single exon probe; asthma; lung cancer; COPD; ILD;

KW chronic obstructive pulmonary disease; interstitial lung disease;

KW familial idiopathic pulmonary fibrosis; neurofibromatosis;

KW tuberous sclerosis; Gaucher's disease; Niemann-pick disease;

KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;

KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;

KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;

KW primary ciliary dyskinesia; pulmonary hypertension;

KW hyaline membrane disease.

XX Homo sapiens.

OS

XX WO2000186003-A2.

XX 15-NOV-2001.

XX 30-JAN-2001; 2001WO-US00665.

XX 04-FEB-2000; 2000US-180312P.

XX 26-MAY-2000; 2000US-207456P.

XX 30-JUN-2000; 2000US-0608408.

XX 03-AUG-2000; 2000US-0632366.

XX 21-SEP-2000; 2000US-234687P.

XX 27-SEP-2000; 2000US-236359P.

XX 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2002-114183/15.

XX Spatially-addressable set of single exon nucleic acid probes, used to

XX measure gene expression in human lung samples -

XX Claim 27; SEQ ID No 35381; 634pp; English.

XX The invention relates to a spatially-addressable set of single exon

XX nucleic acid probes for measuring gene expression in a sample derived

XX from human lung comprising single exon nucleic acid probes having one of

XX 12614 nucleic acid sequences mentioned in the specification, or their

XX complements or the 12387 open reading frames derived from the 12614

XX probes. Also included are a microarray comprising the novel set of

XX nucleic acid expressed in the human lung; measuring gene expression to a

XX sample derived from human lung, comprising (a) contacting the array with

XX a collection of detectably labeled nucleic acids derived from human lung

XX mRNA, and (b) measuring the label detectably bound to each probe of

XX the array; identifying exons in a eukaryotic genome, comprising

XX (a) algorithmically predicting at least one exon from genomic sequences

CC of the eukaryote; and (b) detecting specific hybridisation of detectably
CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
CC having a fragment identical to the predicted exon, the probe is included
CC in the above mentioned microarray; assigning exons to a single gene,
CC comprising (a) identifying exons from genomic sequence by the method
CC above and (b) measuring the expression of each of the exons in several
CC tissues and/or cell types using hybridisation to a single exon
CC microarrays having a probe with the exon, where a common pattern of
CC expression of the exons in the tissues and/or cell types indicates that
CC the exons should be assigned to a single gene; a peptide comprising one
CC of 12011 sequences, mentioned in the specification, or encoded by the
CC probes/open reading frames (ORF). The probes are used for gene
CC expression analysis, and for identifying exons in a gene, particularly
CC using human lung derived mRNA and for the study of lung diseases
CC such as asthma, lung cancer, chronic obstructive pulmonary disease
CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
CC fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,
CC Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary
CC haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis,
CC pulmonary alveolar proteinosis, Karagener syndrome, fibrocystic
CC pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension
CC and hyaline membrane disease. The present sequence is a peptide/protein
CC encoded by a single exon probe of the invention.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 69 AA;

Query Match 100.0%; Score 23; DB 23; Length 69;

Best Local Similarity 100.0%; Pred. No. 3.5e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4

||||

Db 34 DYDA 37

RESULT 33

AAP81871

ID AAP81871 standard; protein; 70 AA.

XX

AC AAP81871;

XX

DT 25-JAN-1991 (first entry)

XX

DE Sequence encoded by of Trigonopsis variabilis D-amino acid oxidase gene,

XX C-terminal.

XX

KW Trigonopsis variabilis D-amino oxidase; Trigonopsis variabilis CBS 4095;

XX high yields of D-amino acid oxidase; enzyme.

XX

OS Trigonopsis variabilis CBS 4095.

XX

PN JP63071180-A.

XX

PD 31-MAR-1988.

XX

PF 16-SEP-1986; 86JP-0215878.

XX

PR 16-SEP-1986; 86JP-0215878.

XX

PA (ASAH) ASahi CHEMICAL IND KK.

XX

DR WPI; 1988-128942/19.

XX

N-PSDB; AAN81684.

XX

PT D-amino acid oxidase (I) gene - derived from Trigonopsis variabilis

XX

PS Disclosure; Fig 2, pages 551; 12pp; Japanese.

XX

XX Using the DNA, (AAN81684) D-aminoacid oxidase (D-AO) can be prepd. in a

CC

CC high yield. A wide range of applications is possible. The whole DNA was
 CC extd. from T. variabilis CBS 4095 by Cryer's method. 40 ug of the DNA
 CC is reacted with 4 units of the restriction enzyme XhoI at 37 degrees C
 CC for 15 min. The reaction liquor is extracted with phenol/chloroform
 CC (1:1). DNAs are pptd. with ethanol and purified by agarose gel
 CC electrophoresis and electro-elution and dissolved in 20 ul of 0.1 fold
 CC concn. of TE buffer soln. (iii) to give a soln. (iv). 30 ug of vector
 CC pUC 18 is cleaved with Bam HI. The resultant DNA fragment is dissolved
 CC in 20 ul of 0.1 fold concn. of (iii) to give a soln. (v). (iv) and (v)
 CC are mixed in a ratio of 3:1 and reacted with T4-DNA ligase to combine the
 CC DNA fragment of the vector and CBS 4095. A DNA library is prepd. The
 CC amino gp. terminal AA sequence of D-AO is determined. A DNA probe is
 CC synthesized. Candidates of the D-AO-clone are selected and sepd. from the
 CC DNA library. The D-AO-clone is identified and the DNA base sequence is
 CC determined. D-AO gene is modified. A generation vector is prepd. and the
 CC D-AO gene is generated. D-methionine and cephalosporin C are oxidized
 CC with E.coli contg. the recombinant.

SQ Sequence 70 AA;

Query Match 100.0%; Score 23; DB 9; Length 70;
 Best Local Similarity 100.0%; Pred. No. 3.5e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
 ||||
 Db 54 DYDA 57

RESULT 34

AAG10868
 ID AAG10868 standard; Protein; 70 AA.

AC AAG10868;

DT 17-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 9363.

KW Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.

OS Arabidopsis thaliana.

PN EP1033405-A2.

PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

PR 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127482.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 28-APR-1999; 99US-0130891.

PR 11-MAY-1999; 99US-0134256.
 PR 14-MAY-1999; 99US-0134218.
 PR 14-MAY-1999; 99US-0134219.
 PR 14-MAY-1999; 99US-0134221.
 PR 18-MAY-1999; 99US-0134370.
 PR 19-MAY-1999; 99US-0134768.
 PR 20-MAY-1999; 99US-0135124.
 PR 21-MAY-1999; 99US-0135124.
 PR 21-MAY-1999; 99US-0135353.
 PR 24-MAY-1999; 99US-0135629.
 PR 25-MAY-1999; 99US-0136021.
 PR 27-MAY-1999; 99US-0136392.
 PR 28-MAY-1999; 99US-0136782.
 PR 01-JUN-1999; 99US-0137222.
 PR 03-JUN-1999; 99US-0137528.
 PR 04-JUN-1999; 99US-0137502.
 PR 07-JUN-1999; 99US-0137724.
 PR 08-JUN-1999; 99US-0138094.
 PR 10-JUN-1999; 99US-0138540.
 PR 10-JUN-1999; 99US-0138847.
 PR 14-JUN-1999; 99US-0139119.
 PR 16-JUN-1999; 99US-0139452.
 PR 17-JUN-1999; 99US-0139453.
 PR 18-JUN-1999; 99US-0139492.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139458.
 PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139460.
 PR 18-JUN-1999; 99US-0139461.
 PR 18-JUN-1999; 99US-0139462.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139464.
 PR 18-JUN-1999; 99US-0139750.
 PR 18-JUN-1999; 99US-0139763.
 PR 21-JUN-1999; 99US-0139817.
 PR 22-JUN-1999; 99US-0139899.
 PR 23-JUN-1999; 99US-0140353.
 PR 23-JUN-1999; 99US-0140354.
 PR 24-JUN-1999; 99US-0140695.
 PR 28-JUN-1999; 99US-0140823.
 PR 29-JUN-1999; 99US-0140991.
 PR 30-JUN-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141842.
 PR 02-JUL-1999; 99US-0142055.
 PR 06-JUL-1999; 99US-0142390.
 PR 08-JUL-1999; 99US-0142803.
 PR 09-JUL-1999; 99US-0142920.
 PR 12-JUL-1999; 99US-0142977.
 PR 13-JUL-1999; 99US-0143542.
 PR 14-JUL-1999; 99US-0143624.
 PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.
 PR 16-JUL-1999; 99US-0144086.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.
 PR 19-JUL-1999; 99US-0144334.
 PR 19-JUL-1999; 99US-0144335.
 PR 20-JUL-1999; 99US-0144352.
 PR 20-JUL-1999; 99US-0144632.
 PR 20-JUL-1999; 99US-0144884.
 PR 21-JUL-1999; 99US-0144814.
 PR 21-JUL-1999; 99US-0145086.
 PR 21-JUL-1999; 99US-0145088.
 PR 22-JUL-1999; 99US-0145085.
 PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145089.
 PR 22-JUL-1999; 99US-0145192.
 PR 23-JUL-1999; 99US-0145145.

PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.

PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161982.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 100.0%; Score 23; DB 21; Length 70;

Best Local Similarity 100.0%; Pred. No. 3.5e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
||||
Db 57 DYDA 60

RESULT 35

AAG54077
ID AAG54077 standard; Protein; 70 AA.

XX AC
AAG54077;

XX DT 18-OCT-2000 (first entry)

XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 68914.

XX KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.

XX OS Arabidopsis thaliana.

XX PN EP1033405-A2.

XX PD 06-SEP-2000.

XX 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 23-APR-1999; 99US-0130891.

PR 28-APR-1999; 99US-0131449.

PR 30-APR-1999; 99US-0132048.

PR 04-MAY-1999; 99US-0132484.

PR 05-MAY-1999; 99US-0132485.

PR 06-MAY-1999; 99US-0132486.

PR 06-MAY-1999; 99US-0132487.

PR 07-MAY-1999; 99US-0132863.

PR 11-MAY-1999; 99US-0134256.

PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134219.

PR 14-MAY-1999; 99US-0134221.

PR 14-MAY-1999; 99US-0134370.

PR 18-MAY-1999; 99US-0134768.

PR 19-MAY-1999; 99US-0134941.

PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137502.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 21-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.

```
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 100.0%; Score 23; DB 21; Length 70;
Best Local Similarity 100.0%; Pred. NO. 3.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
Db 57 DYDA 60

RESULT 36
ABB95902
ID ABB95902 standard; Protein; 79 AA.
AC ABB95902;
XX
DT 21-JUN-2002 (first entry)
XX
DE Human testicular antigen SEQ ID NO: 1286.
XX
KW Human; testicular antigen; testes; cancer; metastasis; immune disorder;
KW reproductive system disorder; urinary system disorder; gene therapy;
KW cardiovascular disorder; respiratory disorder; neurological disorder;
KW gastrointestinal disease; infection; cytostatic.
XX
OS Homo sapiens.
XX
PN WO200155317-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01329.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216680.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 17-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
```

```
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 11-DEC-2000; 2000US-0251990.
PR 05-JAN-2001; 2001US-0259678.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM,
XX
XX WPI; 2001-483232/52.
XX
XX Nucleic acids encoding 973 human testicular antigen polypeptides,
XX useful for preventing, diagnosing and/or treating testicular cancer -
XX
XX Claim 11; SEQ ID NO 1286; 766pp; English.
XX
XX The present invention provides the protein and coding sequences of 973
XX human testicular antigens, and fragments of their genomic sequences. The
XX sequences can be used in the treatment of cardiovascular, urinary system,
XX reproductive system, immune, respiratory, neurological and
XX gastrointestinal disorders, infections, and particularly cancer,
XX especially testicular cancers. The present sequence is a protein of the
XX invention.
XX
XX Sequence 79 AA;
XX
XX Query Match 100.0%; Score 23; DB 22; Length 79;
XX Best Local Similarity 100.0%; Pred. No. 4e+02;
XX Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 DYDA 4
XX
XX Db 8 DYDA 11
XX
XX
XX
XX RESULT 37
XX AAM95203
XX ID AAM95203 standard; Protein; 79 AA.
XX
XX AC AAM95203;
XX
XX XX
XX DT 21-NOV-2001 (first entry)
XX
XX DE Human reproductive system related antigen SEQ ID NO: 3861.
XX
XX KW Human; reproductive system related antigen; reproductive system disorder;
XX KW cancer; gene therapy.
XX
XX OS Homo sapiens.
XX
XX WO200155320-A2.
XX
XX 02-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US01339.
XX
XX 31-JAN-2000; 2000US-0179065.
XX 04-FEB-2000; 2000US-0180628.
XX 24-FEB-2000; 2000US-0184664.
XX 02-MAR-2000; 2000US-0186350.
XX 16-MAR-2000; 2000US-0189874.
XX 17-MAR-2000; 2000US-0190076.
XX 18-APR-2000; 2000US-0198123.
XX 19-MAY-2000; 2000US-0205515.
XX 07-JUN-2000; 2000US-0209467.
XX 28-JUN-2000; 2000US-0214886.
XX 30-JUN-2000; 2000US-0215135.
XX 07-JUL-2000; 2000US-0216647.
XX 07-JUL-2000; 2000US-0216880.
XX 11-JUL-2000; 2000US-0217487.
XX 11-JUL-2000; 2000US-0217496.
XX 14-JUL-2000; 2000US-0218290.
XX 26-JUL-2000; 2000US-0220963.
XX 26-JUL-2000; 2000US-0220964.
XX 14-AUG-2000; 2000US-0224518.
XX 14-AUG-2000; 2000US-0224519.
XX 14-AUG-2000; 2000US-0225213.
XX 14-AUG-2000; 2000US-0225214.
XX 14-AUG-2000; 2000US-0225266.
XX 14-AUG-2000; 2000US-0225267.
XX 14-AUG-2000; 2000US-0225268.
XX 14-AUG-2000; 2000US-0225270.
XX 14-AUG-2000; 2000US-0225447.
XX 14-AUG-2000; 2000US-0225757.
XX 14-AUG-2000; 2000US-0225758.
XX 14-AUG-2000; 2000US-0225759.
XX 18-AUG-2000; 2000US-0226279.
XX 22-AUG-2000; 2000US-0226681.
XX 22-AUG-2000; 2000US-0226688.
XX 22-AUG-2000; 2000US-0227182.
XX 23-AUG-2000; 2000US-0227009.
XX 30-AUG-2000; 2000US-0228924.
XX 01-SEP-2000; 2000US-0229287.
XX 01-SEP-2000; 2000US-0229343.
XX 01-SEP-2000; 2000US-0229344.
XX 01-SEP-2000; 2000US-0229345.
XX 05-SEP-2000; 2000US-0229509.
XX 05-SEP-2000; 2000US-0229513.
XX 06-SEP-2000; 2000US-0230437.
XX 06-SEP-2000; 2000US-0230438.
XX 08-SEP-2000; 2000US-0231242.
XX 08-SEP-2000; 2000US-0231243.
XX 08-SEP-2000; 2000US-0231244.
XX 08-SEP-2000; 2000US-0231413.
XX 08-SEP-2000; 2000US-0231414.
XX 08-SEP-2000; 2000US-0232080.
XX 12-SEP-2000; 2000US-0232081.
XX 12-SEP-2000; 2000US-0231968.
XX 14-SEP-2000; 2000US-0232397.
XX 14-SEP-2000; 2000US-0232398.
XX 14-SEP-2000; 2000US-0232399.
XX 14-SEP-2000; 2000US-0232400.
XX 14-SEP-2000; 2000US-0232401.
XX 14-SEP-2000; 2000US-0233063.
XX 14-SEP-2000; 2000US-0233064.
XX 21-SEP-2000; 2000US-0233065.
XX 21-SEP-2000; 2000US-0234223.
XX 21-SEP-2000; 2000US-0234274.
XX 25-SEP-2000; 2000US-0234597.
XX 25-SEP-2000; 2000US-0234598.
XX 26-SEP-2000; 2000US-0235484.
XX 27-SEP-2000; 2000US-0235834.
```

PR 27-SEP-2000; 2000US-0235836.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.
 PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246529.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 11-DEC-2000; 2000US-0251990.
 PR 05-JAN-2001; 2001US-02559678.
 (HUMA-) HUMAN GENOME SCI INC.

PI Rosen CA, Barash SC, Ruben SM;
 XX WPI: 2001-465570/50.
 DR N-PSDB; AAL01173.
 XX
 PT Isolated nucleic acid molecule encoding a reproductive system antigen -
 PT is used in preventing, treating or ameliorating a medical condition -
 PS Claim 11; SEQ ID NO 3861; 1297pp + Sequence Listing; English.
 XX
 CC The present invention provides the protein and coding sequences of a
 CC number of human reproductive system related antigens. These can be used
 CC in the prevention and treatment of reproductive system disorders,
 CC including cancer. The present sequence is a protein of the invention.
 XX
 SQ Sequence 79 AA;
 Query Match 100.0%; Score 23; DB 22; Length 79;
 Best Local Similarity 100.0%; Pred. No. 4e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DYDA 4
 Db 8 DYDA 11
 ||||
 RESULT 38
 ABP38575
 ID ABP38575 standard; Protein; 79 AA.
 XX
 AC ABP38575;
 XX
 DT 24-JUL-2002 (first entry)
 XX
 DE Staphylococcus epidermidis ORF amino acid sequence SEQ ID NO:3420.
 XX
 KW Staphylococcus epidermidis; open reading frame; ORF; bacterial infection;
 KW antibacterial; gene therapy.
 XX
 OS Staphylococcus epidermidis.
 XX
 PN US6380370-B1.
 XX
 PD 30-APR-2002.
 XX
 PF 13-AUG-1998; 98US-0134001.
 PR 14-AUG-1997; 97US-055779P.
 PR 08-NOV-1997; 97US-064964P.
 XX
 PA (GENO-) GENOME THERAPEUTICS CORP.
 XX
 PI Doucette-Stamm LA, Bush D;
 XX
 WPI: 2002-381255/41.
 DR N-PSDB; ABN91120.
 XX
 PT Novel isolated nucleic acid encoding a Staphylococcus epidermidis
 PT polypeptide, useful for diagnosing and treating bacterial infections -
 XX
 PS Disclosure; SEQ ID 3420; 267pp; English.
 XX
 CC ABN90538 to ABN93374 represent Staphylococcus epidermidis open reading
 CC frame (ORF) nucleic acid sequences which encode the amino acid sequences
 CC given in ABP35124 to ABP37960. The S. epidermidis sequences have
 CC antibacterial activity and can be used in gene therapy. The sequences
 CC can also be used in the diagnosis and treatment of bacterial infections,
 CC particularly S. epidermidis infections. The sequences can be used to
 CC screen for compounds able to interfere with the S. epidermidis life
 CC cycle or inhibit S. epidermidis infection.
 CC N.B. The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from the
 CC USPTO web site.

```
XX SQ Sequence 79 AA;
Query Match 100.0%; Score 23; DB 23; Length 79;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
    ||||
Db 47 DYDA 50

RESULT 39
AAW20370
ID AAW20370 standard; Protein; 81 AA.
XX AC AAW20370;
XX DT 14-JUL-1997 (first entry)
XX DE H. pylori transporter protein 289711.aa.
XX KW Cytoplasmic; vaccine; prevention; treatment; infection; identification;
KW binding compound; bacterium; life cycle; activator; bacteria; inhibitor;
KW duodenal ulcer disease; chronic gastritis; diagnosis; envelope.
XX OS Helicobacter pylori.
XX PN W09640893-A1.
XX PD 19-DEC-1996.
XX PF 06-JUN-1996; 96WO-US09122.
XX PR 01-APR-1996; 96US-0630405.
XX PR 07-JUN-1995; 95US-0487032.
XX PA (ASTR ) ASTRA AB.
XX PI Berglindh OT, Smith D, Mellgaerd BL;
XX WPI; 1997-052306/05.
XX DR N-PSDB; AAT67562.
XX PT Helicobacter pylori nucleic acid; sequences and related
PT polypeptide(s) - useful for vaccines to treat or prevent H. pylori
PT infection, and to detect Helicobacter
XX Claim 56; Page 557; 1481pp; English.
XX This sequence is a H. pylori transporter protein.
XX The protein may be used in a vaccine to prevent or treat H. pylori
XX infection or to identify H. pylori polypeptide binding compounds,
XX useful as potential H. pylori life cycle activators or inhibitors.
XX The genomic sequence of H. pylori (ATCC 55679) was determined from
XX overlapping contigs generated by mechanically shearing the bacterial
XX DNA. The sequences were analysed for ORF of at least 180 nucleotides,
XX and the predicted coding regions defined by computer evaluation. To
XX identify likely H. pylori antigens for vaccine development, the amino
XX acid sequences predicted from various ORF were analysed for significant
XX homology to other known or exported membrane proteins. Having identified
XX and determined the sequences of interest, particular regions can be
XX isolated from H. pylori by PCR amplification for recombinant polypeptide
XX production, e.g. in E. coli hosts.

XX SQ Sequence 81 AA;
Query Match 100.0%; Score 23; DB 18; Length 81;
Best Local Similarity 100.0%; Pred. No. 4.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
    ||||
Db 47 DYDA 50
```

```
Db 41 DYDA 44

RESULT 40
AAG19542
ID AAG19542 standard; Protein; 81 AA.
XX AC AAG19542;
XX DT 17-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 21383.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX OS Arabidopsis thaliana.
XX PN EPI033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-0301439.
XX PR 25-FEB-1999; 99US-0121825.
XX PR 05-MAR-1999; 99US-0123180.
XX PR 09-MAR-1999; 99US-0123548.
XX PR 23-MAR-1999; 99US-0125788.
XX PR 25-MAR-1999; 99US-0126264.
XX PR 29-MAR-1999; 99US-0126785.
XX PR 01-APR-1999; 99US-0127462.
XX PR 06-APR-1999; 99US-0128234.
XX PR 08-APR-1999; 99US-0128714.
XX PR 16-APR-1999; 99US-0129845.
XX PR 19-APR-1999; 99US-0130077.
XX PR 21-APR-1999; 99US-0130449.
XX PR 23-APR-1999; 99US-0130510.
XX PR 28-APR-1999; 99US-0130891.
XX PR 30-APR-1999; 99US-0131449.
XX PR 30-APR-1999; 99US-0132048.
XX PR 04-MAY-1999; 99US-0132407.
XX PR 05-MAY-1999; 99US-0132484.
XX PR 06-MAY-1999; 99US-0132485.
XX PR 06-MAY-1999; 99US-0132486.
XX PR 07-MAY-1999; 99US-0132487.
XX PR 07-MAY-1999; 99US-0132863.
XX PR 11-MAY-1999; 99US-0134256.
XX PR 14-MAY-1999; 99US-0134218.
XX PR 14-MAY-1999; 99US-0134219.
XX PR 14-MAY-1999; 99US-0134221.
XX PR 18-MAY-1999; 99US-0134370.
XX PR 19-MAY-1999; 99US-0134768.
XX PR 20-MAY-1999; 99US-0134941.
XX PR 21-MAY-1999; 99US-0135124.
XX PR 24-MAY-1999; 99US-0135353.
XX PR 25-MAY-1999; 99US-0135629.
XX PR 27-MAY-1999; 99US-0136021.
XX PR 28-MAY-1999; 99US-0136392.
XX PR 01-JUN-1999; 99US-0136782.
XX PR 03-JUN-1999; 99US-0137222.
XX PR 03-JUN-1999; 99US-0137528.
XX PR 07-JUN-1999; 99US-0137724.
XX PR 08-JUN-1999; 99US-0138094.
XX PR 10-JUN-1999; 99US-0138540.
XX PR 14-JUN-1999; 99US-0138847.
XX PR 16-JUN-1999; 99US-0139119.
XX PR 16-JUN-1999; 99US-0139452.
XX PR 17-JUN-1999; 99US-0139453.
XX PR 18-JUN-1999; 99US-0139492.
XX PR 18-JUN-1999; 99US-0139454.
XX PR 18-JUN-1999; 99US-0139455.
```


AC ABG26683;
XX
DT 18-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #26674.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR N-PSDB; AAS90870.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
PS Claim 20; SEQ ID No 57042; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences' (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
SQ Sequence 81 AA;

Query Match 100.0%; Score 23; DB 22; Length 81;
Best Local Similarity 100.0%; Pred. No. 4.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
|||||
Db 68 DYDA 71

RESULT 42
AAW62245
ID AAW62245 standard; protein; 84 AA.
XX
AC AAW62245;
XX

DT 23-SEP-1998 (first entry)
XX
DE Subtilase SEEPiP fragment.
XX
KW Subtilase; variant; mutant; subtilisin; serine protease; Bacillus sp.;
KW auto-proteolytic stability; laundry; dishwashing; proteolytic activity;
KW detergent.
XX
OS Staphylococcus epidermidis.
XX
XX Key Location/Qualifiers
FH Misc-difference 6 /label= unspecified
FT Misc-difference 7 /label= unspecified
FT Misc-difference 8 /label= unspecified
FT Misc-difference 9 /label= unspecified
FT Misc-difference 10 /label= unspecified
FT Misc-difference 11 /label= unspecified
FT Misc-difference 12 /label= unspecified
FT Misc-difference 13 /label= unspecified
FT Misc-difference 14 /label= unspecified
FT Misc-difference 36 /label= unspecified
FT /label= unknown
FT /note= "X represents a gap in the sequence corresponding
FT to residues 146-155 when in alignment with BASBPB"
FT
FT Misc-difference 52 /label= unspecified
FT Misc-difference 53 /label= unspecified
FT Misc-difference 54 /label= unspecified
FT Misc-difference 55 /label= unspecified
FT Misc-difference 56 /label= unspecified
FT Misc-difference 70 /label= unspecified
FT /label= unknown
FT /note= "X represents a gap in the sequence corresponding
FT to residues 173-188 when in alignment with BASBPB"
PN WO9820116-A1.
XX
PD 14-MAY-1998.
XX
PF 04-NOV-1997; 97WO-DK00500.
XX
PR 14-MAR-1997; 97DK-0000284.
PR 04-NOV-1996; 96DK-0001235.
PR 05-NOV-1996; 96DK-0001240.
XX
XX (NOVO) NOVO-NORDISK AS.
PA
XX
XX Andersen C, Bauditz P, Halkier T, Hansen PK, Von Der Osten C;
PI WPI; 1998-286932/25.
XX
XX New subtilase variants with increased auto-proteolytic stability and
XX related nucleic acid, vectors and transformed cells - useful in
XX laundry and dishwashing compositions, having increased proteolytic
XX activity
XX
XX Disclosure; Fig 1; 99pp; English.
XX
XX New subtilase enzyme variants have been derived from a precursor
CC subtilase having either (i) an autoproteolytic cleavage site between


```

DR WPI: 2002-147878/19.
DR N-PSDB; ABQ55428.
XX
PT Isolated nucleic acid molecules encoding novel ovarian polypeptides,
PT useful in the prevention, treatment and diagnosis of cancer (e.g.
PT ovarian cancer), immune disorders, cardiovascular disorders and
PT neurological diseases -
XX
XX Claim 11; SEQ ID NO 3483; 2922pp; English.
XX
CC The invention relates to 2175 novel human ovarian antigens (ABP41054-
CC ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also
CC encompasses polypeptides 90% identical and polynucleotides 95% identical
CC to the sequences of the invention. The invention additionally relates to
CC recombinant vectors and host cells comprising human ovarian antigen
CC polynucleotides, antibodies against human ovarian antigens, and the use
CC of ovarian antigen polynucleotides and polypeptides in diagnosing,
CC treating, prognosing or preventing various ovary and/or breast-related
CC disorders. Such conditions include ovarian cancer and breast cancer, and
CC metastatic tumours of ovarian or breast origin, reproductive system
CC disorders (e.g., infertility, disorders of pregnancy, anovulation,
CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine
CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic
CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and
CC vaginitis), immune disorders (e.g., congenital and acquired
CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),
CC blood-related disorders (e.g., anaemia), cardiovascular disorders,
CC respiratory disorders, neurological disorders, gastrointestinal disorders
CC and urinary system disorders. Ovarian antigen polypeptides and
CC polynucleotides may also be used in screening for compounds which
CC modulate ovarian antigen expression or activity. The polynucleotides may
CC further be used for gene therapy, chromosome mapping, in the
CC identification of individuals and in forensic analysis, and the
CC polypeptides may be used as food additives or to prepare antibodies
CC useful in disease diagnosis, drug targeting and phenotyping. The present
CC sequence represents a human ovarian antigen of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 87 AA;
Query Match 100.0%; Score 23; DB 23; Length 87;
Best Local Similarity 100.0%; Pred. No. 4.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DYDA 4
|
|
|
|
Db 74 DYDA 77
|
|
|
|
RESULT 45
AAG61460
ID AAG61460 standard; Protein; 91 AA.
XX
AC AAG61460;
XX
XX 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 79718.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
XX 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
25-FEB-1999; 99US-0121825.
05-MAR-1999; 99US-0123180.
09-MAR-1999; 99US-0123548.
23-MAR-1999; 99US-0125788.
23-MAR-1999; 99US-0126264.
29-MAR-1999; 99US-0126785.
01-APR-1999; 99US-0127462.
06-APR-1999; 99US-0128234.
08-APR-1999; 99US-0128714.
16-APR-1999; 99US-0129845.
19-APR-1999; 99US-0130077.
21-APR-1999; 99US-0130449.
23-APR-1999; 99US-0130510.
23-APR-1999; 99US-0130891.
28-APR-1999; 99US-0131449.
30-APR-1999; 99US-0132048.
30-APR-1999; 99US-0132407.
04-MAY-1999; 99US-0132484.
05-MAY-1999; 99US-0132485.
06-MAY-1999; 99US-0132486.
06-MAY-1999; 99US-0132487.
07-MAY-1999; 99US-0132863.
11-MAY-1999; 99US-0134256.
14-MAY-1999; 99US-0134218.
14-MAY-1999; 99US-0134219.
14-MAY-1999; 99US-0134221.
14-MAY-1999; 99US-0134370.
18-MAY-1999; 99US-0134768.
19-MAY-1999; 99US-0134941.
20-MAY-1999; 99US-0135124.
21-MAY-1999; 99US-0135353.
24-MAY-1999; 99US-0135629.
25-MAY-1999; 99US-0136021.
27-MAY-1999; 99US-0136392.
28-MAY-1999; 99US-0136782.
01-JUN-1999; 99US-0137222.
03-JUN-1999; 99US-0137528.
04-JUN-1999; 99US-0137502.
07-JUN-1999; 99US-0137724.
08-JUN-1999; 99US-0138094.
10-JUN-1999; 99US-0138540.
10-JUN-1999; 99US-0138847.
14-JUN-1999; 99US-0139119.
16-JUN-1999; 99US-0139452.
16-JUN-1999; 99US-0139453.
17-JUN-1999; 99US-0139492.
18-JUN-1999; 99US-0139454.
18-JUN-1999; 99US-0139455.
18-JUN-1999; 99US-0139456.
18-JUN-1999; 99US-0139457.
18-JUN-1999; 99US-0139458.
18-JUN-1999; 99US-0139459.
18-JUN-1999; 99US-0139460.
18-JUN-1999; 99US-0139461.
18-JUN-1999; 99US-0139462.
18-JUN-1999; 99US-0139463.
18-JUN-1999; 99US-0139750.
18-JUN-1999; 99US-0139763.
21-JUN-1999; 99US-0139817.
22-JUN-1999; 99US-0139899.
23-JUN-1999; 99US-0140353.
23-JUN-1999; 99US-0140354.
24-JUN-1999; 99US-0140695.
28-JUN-1999; 99US-0140823.
28-JUN-1999; 99US-0140991.
30-JUN-1999; 99US-0141287.
01-JUL-1999; 99US-0141842.
01-JUL-1999; 99US-0142154.
02-JUL-1999; 99US-0142055.
06-JUL-1999; 99US-0142390.
08-JUL-1999; 99US-0142803.
09-JUL-1999; 99US-0142920.
12-JUL-1999; 99US-0142977.

```

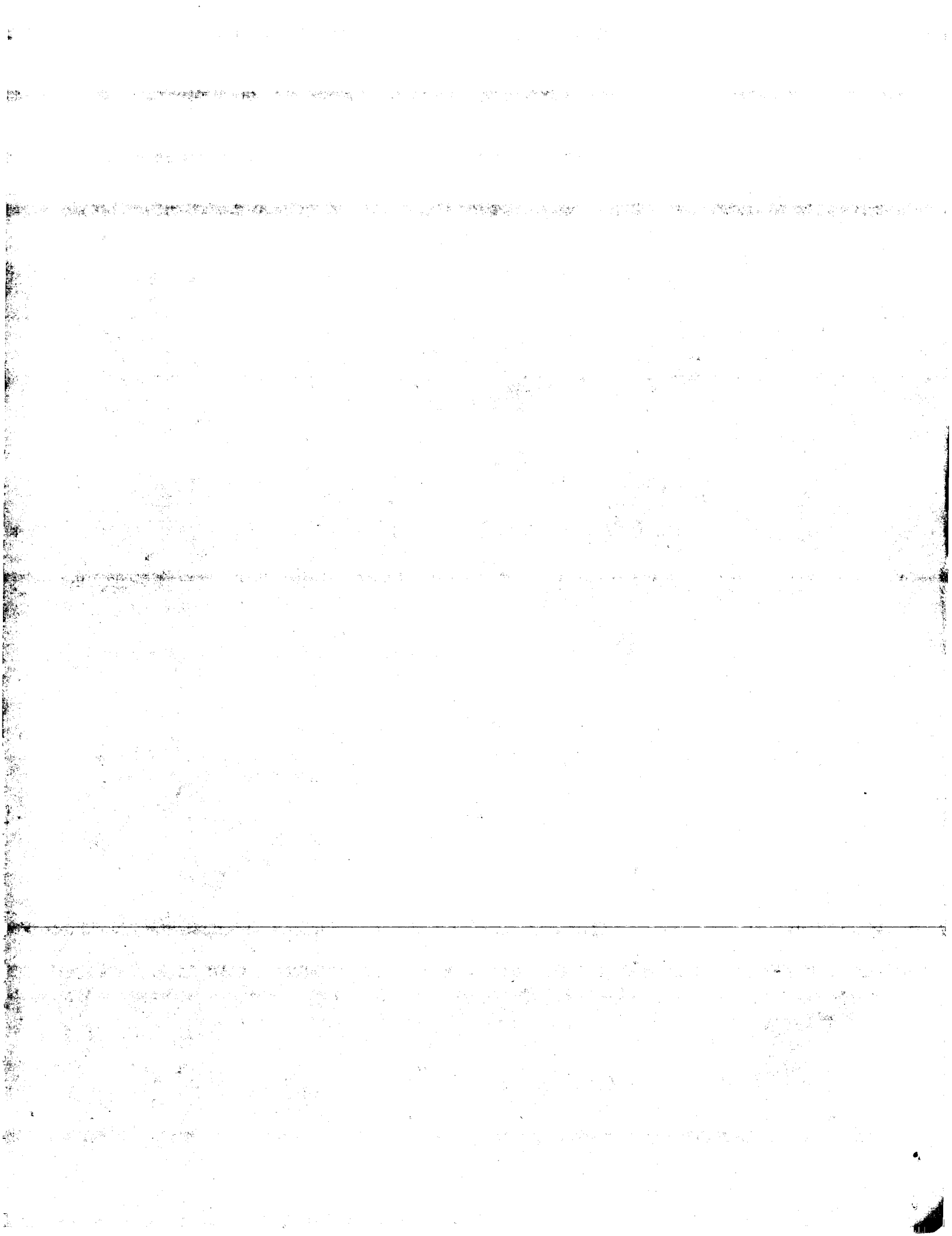
PR 13-JUL-1999; 99US-0143542.
 PR 14-JUL-1999; 99US-0143624.
 PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.
 PR 16-JUL-1999; 99US-0144086.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.
 PR 19-JUL-1999; 99US-0144334.
 PR 19-JUL-1999; 99US-0144335.
 PR 20-JUL-1999; 99US-0144352.
 PR 20-JUL-1999; 99US-0144632.
 PR 20-JUL-1999; 99US-0144884.
 PR 21-JUL-1999; 99US-0144814.
 PR 21-JUL-1999; 99US-0145086.
 PR 21-JUL-1999; 99US-0145088.
 PR 22-JUL-1999; 99US-0145085.
 PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145089.
 PR 22-JUL-1999; 99US-0145192.
 PR 23-JUL-1999; 99US-0145145.
 PR 23-JUL-1999; 99US-0145218.
 PR 23-JUL-1999; 99US-0145224.
 PR 26-JUL-1999; 99US-0145276.
 PR 27-JUL-1999; 99US-0145913.
 PR 27-JUL-1999; 99US-0145918.
 PR 27-JUL-1999; 99US-0145919.
 PR 28-JUL-1999; 99US-0145951.
 PR 02-AUG-1999; 99US-0146386.
 PR 02-AUG-1999; 99US-0146388.
 PR 02-AUG-1999; 99US-0146389.
 PR 03-AUG-1999; 99US-0147038.
 PR 04-AUG-1999; 99US-0147204.
 PR 04-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147192.
 PR 05-AUG-1999; 99US-0147260.
 PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-0147493.
 PR 09-AUG-1999; 99US-0147935.
 PR 10-AUG-1999; 99US-0148171.
 PR 11-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149368.
 PR 17-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 20-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151438.
 PR 01-SEP-1999; 99US-0151930.
 PR 07-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.

PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 100.0%; Score 23; DB 21; Length 91;
 Best Local Similarity 100.0%; Pred. No. 4.7e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DYDA 4
 Db 78 DYDA 81

Search completed: February 6, 2003, 11:19:35
 Job time : 29.8333 secs



GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 6, 2003, 11:16:04 ; Search time 9.33333 Seconds
(without alignments)
12.610 Million cell updates/sec

Title: PAT943-1
Perfect score: 23
Sequence: 1 dyda 4

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues
Total number of hits satisfying chosen parameters: 175

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 100%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA.*
1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/1aa/PTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	23	100.0	20	1	US-08-442-884-4
2	23	100.0	21	6	5268270-6
3	23	100.0	24	2	US-08-637-759B-313
4	23	100.0	24	3	US-08-871-355A-313
5	23	100.0	24	4	US-09-201-945-313
6	23	100.0	28	2	US-08-392-816-4
7	23	100.0	38	4	US-08-630-915A-102
8	23	100.0	39	4	US-08-630-915A-104
9	23	100.0	40	4	US-08-630-915A-96
10	23	100.0	45	1	US-08-072-070-3
11	23	100.0	45	1	US-08-214-164-3
12	23	100.0	45	1	US-08-469-434-3
13	23	100.0	45	1	US-08-214-222-3
14	23	100.0	45	2	US-08-467-852A-4
15	23	100.0	45	2	US-08-468-718-3
16	23	100.0	45	2	US-08-247-491A-4
17	23	100.0	45	3	US-08-312-949-3
18	23	100.0	45	3	US-08-446-201-1
19	23	100.0	47	4	US-09-314-268-154
20	23	100.0	49	3	US-08-651-136C-58
21	23	100.0	49	4	US-09-229-911A-58
22	23	100.0	64	2	US-08-749-526-3
23	23	100.0	66	1	US-08-140-797-8
24	23	100.0	66	1	US-08-486-670A-8
25	23	100.0	66	1	US-08-241-853-7
26	23	100.0	66	2	US-08-850-917-7
27	23	100.0	66	2	US-08-749-526-2

Sequence 15, Appl
Sequence 3420, Ap
Sequence 7, Appli
Sequence 4, Appli
Sequence 10, Appl
Sequence 5, Appli
Sequence 20, Appl
Sequence 5, Appli
Sequence 509, App
Sequence 53, Appl
Sequence 1, Appli
Sequence 92, Appl
Sequence 92, Appl
Sequence 92, Appl
Sequence 4, Appli
Sequence 4, Appli
Sequence 4, Appli
Sequence 4, Appli

ALIGNMENTS

RESULT 1
US-08-442-884-4
; Sequence 4, Application US/08442884
; Patent No. 5637490
; GENERAL INFORMATION:
; APPLICANT: Mutsumi SANO et al.
; TITLE OF INVENTION: Alpha-1,3/4-FUCOSIDASE GENE
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005

COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/442.884
; FILING DATE: May 17, 1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek, Jr.
; REGISTRATION NUMBER: 33,367
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-8850
; TELEFAX:
; TELEX:

INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal fragment
US-08-442-884-4

Query Match 100.0%; Score 23; DB 1; Length 20;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4

Db 10 DYDA 13
||||

RESULT 2
; Patent No. 5268270
; APPLICANT: Meyer, Thomas F.; Halter, Roman; Pohlner, Johannes
; TITLE OF INVENTION: PROCESS FOR PRODUCING PROTEINS USING GRAM
; NEGATIVE HOST CELLS
; NUMBER OF SEQUENCES: 6
; CURRENT APPLICATION NUMBER: US/07/171,872
; FILING DATE: 01-JUL-1987
; SEQ ID NO: 6:
; LENGTH: 21
5268270-6

Query Match 100.0%; Score 23; DB 6; Length 21;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
||||
Db 16 DYDA 19

RESULT 3
US-08-637-759B-313
; Sequence 313, Application US/08637759B
; Patent No. 5876931
; GENERAL INFORMATION:
; APPLICANT: David William Holden
; TITLE OF INVENTION: Identification of Genes
; NUMBER OF SEQUENCES: 501
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 2800 One Atlantic Center
; STREET: 1201 West Peachtree Street
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30309-3450
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/637,759B
; FILING DATE: 03-MAY-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB95/02875
; FILING DATE: 11-DEC-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: RPMS 101
; TELEPHONE: (404) 873-8794
; TELEFAX: (404) 873-8795
; INFORMATION FOR SEQ ID NO: 313:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 24 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
US-08-637-759B-313

Query Match 100.0%; Score 23; DB 2; Length 24;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
||||
Db 19 DYDA 22

RESULT 4
US-08-871-355A-313
; Sequence 313, Application US/08871355A
; Patent No. 6015669
; GENERAL INFORMATION:
; APPLICANT: David William Holden
; TITLE OF INVENTION: Identification of Genes
; NUMBER OF SEQUENCES: 501
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 2800 One Atlantic Center
; STREET: 1201 West Peachtree Street
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30309-3450
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/871,355A
; FILING DATE: 09-JUN-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB95/02875
; FILING DATE: 11-DEC-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: RPMS 101 CON
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 873-8794
; TELEFAX: (404) 873-8795
; INFORMATION FOR SEQ ID NO: 313:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 24 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
US-08-871-355A-313

Query Match 100.0%; Score 23; DB 3; Length 24;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
||||
Db 19 DYDA 22

RESULT 5
US-09-201-945-313
; Sequence 313, Application US/09201945
; Patent No. 6342215
; GENERAL INFORMATION:
; APPLICANT: David William Holden
; TITLE OF INVENTION: Identification of Genes
; NUMBER OF SEQUENCES: 501
; CORRESPONDENCE ADDRESS:

ADDRESSEE: Patrea L. Pabst
STREET: 2800 One Atlantic Center
CITY: Atlanta
STATE: Georgia
COUNTRY: USA
ZIP: 30309-3450
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/201,945
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/637,759
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: RPMS 101
TELECOMMUNICATION INFORMATION:
TELEPHONE: (404) 873-8794
TELEFAX: (404) 873-8795
INFORMATION FOR SEQ ID NO: 313:
SEQUENCE CHARACTERISTICS:
LENGTH: 24 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
US-09-201-945-313

Query Match 100.0%; Score 23; DB 4; Length 24;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
DB 19 DYDA 22

RESULT 6
US-08-392-816-4
Sequence 4, Application US/08392816
Patent No. 5859333
GENERAL INFORMATION:
APPLICANT: Keeling, Peter L.
APPLICANT: Lomako, Joseph
APPLICANT: Gleowar-Singh, Dave
APPLICANT: Singletary, George W.
APPLICANT: Whelan, William J.
TITLE OF INVENTION: NOVEL PLANTS AND PROCESSES FOR MAKING THEM
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pillsbury Madison & Sutro, L.L.P.
STREET: 1100 New York Avenue, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3918
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Microsoft Word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/392,816
FILING DATE: 18-DEC-1995

CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB93/01821
FILING DATE: 26-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9218185.8
FILING DATE: 26-AUG-1992
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 28 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-392-816-4

Query Match 100.0%; Score 23; DB 2; Length 28;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
DB 17 DYDA 20

RESULT 7
US-08-630-915A-102
Sequence 102, Application US/08630915A
Patent No. 6309820
GENERAL INFORMATION:
APPLICANT: SPARKS, Andrew B.
APPLICANT: HOFFMAN, No. 6309820h
APPLICANT: KAY, Brian K.
APPLICANT: FOWLES, Dana M.
APPLICANT: MCCONNELL, Stephen J.
TITLE OF INVENTION: POLYPEPTIDES HAVING A FUNCTIONAL
DOMAIN OF INTEREST AND METHODS OF IDENTIFYING AND
TITLE OF INVENTION: USING SAME
NUMBER OF SEQUENCES: 227
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA: US/08/630,915A
APPLICATION NUMBER: US/08-APR-1996
FILING DATE: 03-APR-1996
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Mirock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 1101-174
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 102:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-630-915A-102

Query Match 100.0%; Score 23; DB 4; Length 38;

```
Best Local Similarity 100.0%; Pred. No. 62;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
    ||||
Db 5 DYDA 8

RESULT 8
US-08-630-915A-104
; Sequence 104, Application US/08630915A
; Patent No. 6309820
; GENERAL INFORMATION:
; APPLICANT: SPARKS, Andrew B.
; APPLICANT: HOFFMAN, No. 6309820h
; APPLICANT: KAY, Brian K.
; APPLICANT: FOWLKES, Dana M.
; APPLICANT: MCCONNELL, Stephen J.
; TITLE OF INVENTION: POLYPEPTIDES HAVING A FUNCTIONAL
; TITLE OF INVENTION: DOMAIN OF INTEREST AND METHODS OF IDENTIFYING AND
; NUMBER OF SEQUENCES: 227
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/630,915A
; FILING DATE: 03-APR-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Mistrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 1101-174
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 96:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 40 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; US-08-630-915A-96

Query Match 100.0%; Score 23; DB 4; Length 40;
Best Local Similarity 100.0%; Pred. No. 66;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
    ||||
Db 5 DYDA 8

RESULT 10
US-08-072-070-3
; Sequence 3, Application US/08072070
; Patent No. 5476929
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: Yoether, Janet L
; APPLICANT: McDaniel, Larry S
; TITLE OF INVENTION: STRUCTURAL GENE OF PNEUMOCOCCAL
; TITLE OF INVENTION: PROTEIN
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSES: Shoemaker and Mattare, Ltd
; STREET: Suite 1203, 2001 Jefferson Davis Highway
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202-0286
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/072,070
; FILING DATE: 19930603

Best Local Similarity 100.0%; Pred. No. 62;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
    ||||
Db 5 DYDA 8

RESULT 9
US-08-630-915A-96
; Sequence 96, Application US/08630915A
; Patent No. 6309820
; GENERAL INFORMATION:
; APPLICANT: SPARKS, Andrew B.
; APPLICANT: HOFFMAN, No. 6309820h
; APPLICANT: KAY, Brian K.
; APPLICANT: FOWLKES, Dana M.
```



```
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/835,698
; FILING DATE: 12-FEB-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/656,773
; FILING DATE: 15-FEB-1991
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 415-0810
; TELEFAX: (703) 521-0378
; TELEX: LUKPAT WASHINGTON
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 45 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; US-08-072-070-3

Query Match
Best Local Similarity 100.0%; Score 23; DB 1; Length 45;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
Db 14 DYDA 17

RESULT 11
US-08-214-164-3
; Sequence 3, Application US/08214164
; Patent No. 5728387
; GENERAL INFORMATION:
; APPLICANT: BRILES, DAVID E.
; APPLICANT: YOTHER, JANET L.
; TITLE OF INVENTION: STRUCTURAL GENE OF PNEUMOCOCCAL PROTEIN
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Shoemaker and Mattare, Ltd
; STREET: Suite 1203, 2001 Jefferson Davis Highway
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202-0286
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/214,164
; FILING DATE: 17-MAR-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/656,773
; FILING DATE: 15-FEB-1991
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Berkstresser, Jerry W.
; REGISTRATION NUMBER: 22,651
; REFERENCE/DOCKET NUMBER: 6102-137
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 415-0810
; TELEFAX: (703) 521-0813
; TELEX: LUKPAT WASHINGTON
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 45 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-214-164-3

Query Match
100.0%; Score 23; DB 1; Length 45;
```

```
Best Local Similarity 100.0%; Pred. No. 75;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
Db 14 DYDA 17

RESULT 12
US-08-469-434-3
; Sequence 3, Application US/08469434
; Patent No. 5753463
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: Yother, Janet L
; APPLICANT: McDaniel, Larry S
; TITLE OF INVENTION: STRUCTURAL GENE OF PNEUMOCOCCAL
; TITLE OF INVENTION: PROTEIN
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Shoemaker and Mattare, Ltd
; STREET: Suite 1203, 2001 Jefferson Davis Highway
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202-0286
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/469,434
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/072,065
; FILING DATE: 03 JUNE 1993
; APPLICATION NUMBER: US/07/835,698
; FILING DATE: 12-FEB-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/656,773
; FILING DATE: 15-FEB-1991
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 415-0810
; TELEFAX: (703) 521-0378
; TELEX: LUKPAT WASHINGTON
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 45 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; US-08-469-434-3

Query Match
100.0%; Score 23; DB 1; Length 45;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
Db 14 DYDA 17

RESULT 13
US-08-214-222-3
; Sequence 3, Application US/08214222
; Patent No. 5804193
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: Yother, Janet L
; TITLE OF INVENTION: STRUCTURAL GENE OF PNEUMOCOCCAL
; TITLE OF INVENTION: PROTEIN
; NUMBER OF SEQUENCES: 6
```

;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Shoemaker and Mattare, Ltd
;; STREET: Suite 1203, 2001 Jefferson Davis Highway
;; CITY: Arlington
;; STATE: Virginia
;; COUNTRY: U.S.A.
;; ZIP: 22202-0286
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentin Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/214,222
;; FILING DATE: 17-MAR-1994
;; CLASSIFICATION: 530
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/835,698
;; FILING DATE: 12-FEB-1992
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (703) 415-0810
;; TELEFAX: (703) 521-0378
;; TELEX: LUKPAT WASHINGTON
;; INFORMATION FOR SEQ ID NO: 3:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 45 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; US-08-214-222-3

Query Match 100.0%; Score 23; DB 1; Length 45;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DYDA 4
Db 14 DYDA 17

RESULT 14
US-08-467-852A-4
;; Sequence 4, Application US/08467852A
;; Patent No. 5856170
;; GENERAL INFORMATION:
;; APPLICANT: BRILES, David E.
;; APPLICANT: YOTTER, Janet L.
;; APPLICANT: MCDANIEL, Larry S.
;; TITLE OF INVENTION: STRUCTURAL GENE OF PNEUMOCOCCAL PROTEIN
;; NUMBER OF SEQUENCES: 10
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: FROMMER LAWRENCE & HAUG LLP
;; STREET: 745 Fifth Avenue
;; CITY: New York
;; STATE: NY
;; COUNTRY: USA
;; ZIP: 10151
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentin Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/467,852A
;; FILING DATE: 06-JUN-1995
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: KOWALSKI, Thomas J.
;; REGISTRATION NUMBER: 32,147
;; REFERENCE/DOCKET NUMBER: 454312-2064
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 212-588-0800
;; TELEFAX: 212-588-0500
;; INFORMATION FOR SEQ ID NO: 4:

;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 45 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: n/a
;; TOPOLOGY: linear
;; MOLECULE TYPE: amino acid
;; US-08-467-852A-4

Query Match 100.0%; Score 23; DB 2; Length 45;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DYDA 4
Db 14 DYDA 17

RESULT 15
US-08-468-718-3
;; Sequence 3, Application US/08468718
;; Patent No. 5871943
;; GENERAL INFORMATION:
;; APPLICANT: Briles, David E
;; APPLICANT: Yotter, Janet L
;; APPLICANT: Mcdaniel, Larry S
;; TITLE OF INVENTION: STRUCTURAL GENE OF PNEUMOCOCCAL
;; TITLE OF INVENTION: PROTEIN
;; NUMBER OF SEQUENCES: 6
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Shoemaker and Mattare, Ltd
;; STREET: Suite 1203, 2001 Jefferson Davis Highway
;; CITY: Arlington
;; STATE: Virginia
;; COUNTRY: U.S.A.
;; ZIP: 22202-0286
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentin Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/468,718
;; FILING DATE:
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US/08/072,068
;; FILING DATE: 03 JUNE 1993
;; APPLICATION NUMBER: US/07/835,698
;; FILING DATE: 12-FEB-1992
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US/07/656,773
;; FILING DATE: 15-FEB-1991
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (703) 415-0810
;; TELEFAX: (703) 521-0378
;; TELEX: LUKPAT WASHINGTON
;; INFORMATION FOR SEQ ID NO: 3:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 45 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; US-08-468-718-3

Query Match 100.0%; Score 23; DB 2; Length 45;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DYDA 4
Db 14 DYDA 17

RESULT 16

US-08-247-491A-4
; Sequence 4, Application US/08247491A
; Patent No. 5965400
; GENERAL INFORMATION:
; APPLICANT: BRILES, David E.
; APPLICANT: FOTHER, Janet L.
; TITLE OF INVENTION: STRUCTURAL GENE OF PNEUMOCOCCAL PROTEIN
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FROMMER LAWRENCE & HAUG LLP
; STREET: 745 Fifth Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10151
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/247,491A
; FILING DATE: 23-JUN-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: KOWALSKI, Thomas J.
; REGISTRATION NUMBER: 32,147
; REFERENCE/DOCKET NUMBER: 454312-2041
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-588-0800
; TELEFAX: 212-588-0500
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 45 amino acids
; TYPE: amino acid
; STRANDEDNESS: n/a
; TOPOLOGY: linear
; MOLECULE TYPE: amino acid
US-08-247-491A-4

Query Match 100.0%; Score 23; DB 2; Length 45;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
Db 14 DYDA 17

RESULT 17
US-08-312-949-3
; Sequence 3, Application US/08312949
; Patent No. 6027734
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: Wu, Hong-Yin
; TITLE OF INVENTION: MUCOSAL ADMINISTRATION OF
; PNEUMOCOCCAL ANTIGENS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/312,949

; FILING DATE: 30-SEP-1994
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2049
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 45 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-312-949-3

Query Match 100.0%; Score 23; DB 3; Length 45;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
Db 14 DYDA 17

RESULT 18
US-08-446-201-1
; Sequence 1, Application US/08446201B
; Patent No. 6042838
; GENERAL INFORMATION:
; APPLICANT: BRILES, David E.
; APPLICANT: WU, Hong-Yin
; TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS FOR MUCOSAL ADMINISTRATION OF
; PNEUMOCOCCAL SURFACE PROTEIN A (PSPA)
; FILE REFERENCE: 454312-2018
; CURRENT APPLICATION NUMBER: US/08/446,201B
; CURRENT FILING DATE: 1995-05-19
; EARLIER APPLICATION NUMBER: 08/312,949
; EARLIER FILING DATE: 1994-09-30
; EARLIER APPLICATION NUMBER: 08/246,636
; EARLIER FILING DATE: 1994-05-20
; EARLIER APPLICATION NUMBER: 08/048,896
; EARLIER FILING DATE: 1993-04-20
; EARLIER APPLICATION NUMBER: 07/835,698
; EARLIER FILING DATE: 1992-02-12
; EARLIER APPLICATION NUMBER: 07/656,773
; EARLIER FILING DATE: 1991-02-15
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; TYPE: PRT
; LENGTH: 45
; ORGANISM: Streptococcus pneumoniae
US-08-446-201-1

Query Match 100.0%; Score 23; DB 3; Length 45;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
Db 14 DYDA 17

RESULT 19
US-09-314-268-154
; Sequence 154, Application US/09314268
; Patent No. 6346377
; GENERAL INFORMATION:
; APPLICANT: Doorbar, John
; TITLE OF INVENTION: IMPROVEMENTS IN OR RELATING TO SCREENING FOR PAPILLOMA

;
; TITLE OF INVENTION: VIRUSES
; FILE REFERENCE: 3789/80902
; CURRENT APPLICATION NUMBER: US/09/314,268
; CURRENT FILING DATE: 1999-03-19
; EARLIER APPLICATION NUMBER: 09/314,268
; EARLIER FILING DATE: 1999-05-18
; NUMBER OF SEQ ID NOS: 179
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 154
; LENGTH: 47
; TYPE: PRT
; ORGANISM: Human papillomavirus type 60
US-09-314-268-154

Query Match 100.0%; Score 23; DB 4; Length 47;
Best Local Similarity 100.0%; Pred. No. 78;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
||||
DB 23 DYDA 26

RESULT 20
US-08-651-136C-58
; Sequence 58, Application US/08651136C
; Patent No. 6001639
; GENERAL INFORMATION:
; APPLICANT: Schulein, Martin
; APPLICANT: Andersen, Lene N.
; APPLICANT: Lassen, Soren F.
; APPLICANT: Kauppinen, Markus S.
; APPLICANT: Lange, Lene
; APPLICANT: Nielsen, Ruby I.
; APPLICANT: Ihara, Michiko
; APPLICANT: Takagi, Shinobu
; TITLE OF INVENTION: No. 6001639el Endoglucanases
; NUMBER OF SEQUENCES: 109
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6001639o No. 6001639disk of No. 6001639th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/651,136C
; FILING DATE: 21-MAY-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J.
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 4366,200-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 58:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 49 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-651-136C-58

Query Match 100.0%; Score 23; DB 3; Length 49;
Best Local Similarity 100.0%; Pred. No. 82;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
||||
DB 11 DYDA 14

RESULT 21
US-09-229-911A-58
; Sequence 58, Application US/09229911A
; Patent No. 6387690
; GENERAL INFORMATION:
; APPLICANT: Schulein, Martin
; APPLICANT: Andersen, Lene N.
; APPLICANT: Lassen, Soren F.
; APPLICANT: Kauppinen, Markus S.
; APPLICANT: Lange, Lene
; APPLICANT: Nielsen, Ruby I.
; APPLICANT: Ihara, Michiko
; APPLICANT: Takagi, Shinobu
; TITLE OF INVENTION: No. 6387690el Endoglucanases
; NUMBER OF SEQUENCES: 109
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6387690o No. 6387690disk of No. 6387690th America, I

; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/229,911A
; FILING DATE: 13-Jan-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/651,136
; FILING DATE: 21-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J.
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 4366,200-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 58:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 49 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 58:
US-09-229-911A-58

Query Match 100.0%; Score 23; DB 4; Length 49;
Best Local Similarity 100.0%; Pred. No. 82;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
||||
DB 11 DYDA 14

RESULT 22
US-08-749-526-3
; Sequence 3, Application US/08749526
; Patent No. 5911991
; GENERAL INFORMATION:
; APPLICANT: Pogo, A Oscar
; APPLICANT: Chaudhuri, Asok
; TITLE OF INVENTION: MALARIAL BINDING SITE ON DUFFY
; BLOOD GROUP PROTEIN

```

; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann & Baron, LLP
; STREET: 350 Jericho Turnpike
; CITY: Jericho
; STATE: New York
; COUNTRY: USA
; ZIP: 11753
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
; COMPUTER: IBM compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/749,526
; FILING DATE: 15-NOV-1996
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/140,797
; FILING DATE: 21-OCT-1993
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: O'Dea, Sean W.
; REGISTRATION NUMBER: 37690
; REFERENCE/DOCKET NUMBER: 454-12
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 822-3550
; TELEFAX: (516) 822-3582
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 64 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-749-526-3

Query Match 100.0%; Score 23; DB 2; Length 64;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
Db 40 DYDA 43

RESULT 23
US-08-140-797-8
; Sequence 8, Application US/08140797
; Patent No. 5578714
; GENERAL INFORMATION:
; APPLICANT: Pogo, Angel Oscar; Chaudhuri, Asok
; TITLE OF INVENTION: THE CLONING OF DUFFY BLOOD GROUP ANTIGEN
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sprung Horn Kramer & Woods
; STREET: 660 White Plains Road
; CITY: Tarrytown
; STATE: New York
; COUNTRY: USA
; ZIP: 10591-5144
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 800 kb storage
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: System 7.0
; SOFTWARE: WordPerfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/140,797
; FILING DATE: October 21, 1993
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:

```

```

; NAME: Kurt G. Briscoe
; REGISTRATION NUMBER: 33,141
; REFERENCE/DOCKET NUMBER: NYBC 265-KGB
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (914) 332-1700
; TELEFAX: (914) 332-1844
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 66 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; US-08-140-797-8

Query Match 100.0%; Score 23; DB 1; Length 66;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
Db 42 DYDA 45

RESULT 24
US-08-486-670A-8
; Sequence 8, Application US/08486670A
; Patent No. 5683696
; GENERAL INFORMATION:
; APPLICANT: Pogo, Angel Oscar; Chaudhuri, Asok
; TITLE OF INVENTION: THE CLONING OF DUFFY BLOOD GROUP ANTIGEN
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sprung Horn Kramer & Woods
; STREET: 660 White Plains Road
; CITY: Tarrytown
; STATE: New York
; COUNTRY: USA
; ZIP: 10591-5144
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 800 kb storage
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: System 7.0
; SOFTWARE: WordPerfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,670A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/140,797
; FILING DATE: October 21, 1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Kurt G. Briscoe
; REGISTRATION NUMBER: 33,141
; REFERENCE/DOCKET NUMBER: NYBC 265-KGB
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (914) 332-1700
; TELEFAX: (914) 332-1844
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 66 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; US-08-486-670A-8

Query Match 100.0%; Score 23; DB 1; Length 66;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
Db 42 DYDA 45

RESULT 25

```

US-08-241-853-7
; Sequence 7, Application US/08241853
; Patent No. 5693488
; GENERAL INFORMATION:
; APPLICANT: Fang, Kathy S.
; APPLICANT: Hanafusa, Hidesaburo
; TITLE OF INVENTION: A TRANSMEMBRANE TYROSINE PHOSPHATASE
; TITLE OF INVENTION: AND METHODS OF USE THEREOF
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/241/853
; FILING DATE: 12-MAY-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-078
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 66 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-241-853-7
Query Match 100.0%; Score 23; DB 1; Length 66;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0;
Gaps 0;
QY 1 DYDA 4
Db 1 DYDA 4
RESULT 26
US-08-850-917-7
; Sequence 7, Application US/08850917
; Patent No. 5954045
; GENERAL INFORMATION:
; APPLICANT: Fang, Kathy S.
; APPLICANT: Hanafusa, Hidesaburo
; TITLE OF INVENTION: A TRANSMEMBRANE TYROSINE PHOSPHATASE
; TITLE OF INVENTION: AND METHODS OF USE THEREOF
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/850,917
; FILING DATE: 02-MAY-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/241,853
; FILING DATE: 12-MAY-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-078
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 66 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-850-917-7
Query Match 100.0%; Score 23; DB 2; Length 66;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0;
Gaps 0;
QY 1 DYDA 4
Db 1 DYDA 4
RESULT 27
US-08-749-526-2
; Sequence 2, Application US/08749526
; Patent No. 5911991
; GENERAL INFORMATION:
; APPLICANT: Pogo, A Oscar
; APPLICANT: Chaudhuri, Asok
; TITLE OF INVENTION: MALARIAL BINDING SITE ON DUFFY
; TITLE OF INVENTION: BLOOD GROUP PROTEIN
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann & Baron, LLP
; STREET: 350 Jericho Turnpike
; CITY: Jericho
; STATE: New York
; COUNTRY: USA
; ZIP: 11753
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
; COMPUTER: IBM compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/749,526
; FILING DATE: 15-NOV-1996
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/140,797
; FILING DATE: 21-OCT-1993
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: O'Dea, Sean W.
; REGISTRATION NUMBER: 37690
; REFERENCE/DOCKET NUMBER: 454-12
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 822-3550

TELEFAX: (516) 822-3582
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 66 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 US-08-749-526-2

Query Match 100.0%; Score 23; DB 2; Length 66;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DYDA 4
 Db 42 DYDA 45

RESULT 28
 US-08-963-851-15
 Sequence 15, Application US/08963851
 Patent No. 6300116
 GENERAL INFORMATION:
 APPLICANT: VAN DER OSTEN, CLAUS
 APPLICANT: HALKIER, TORDEN
 APPLICANT: ANDERSEN, CARSTEN
 APPLICANT: BAUDITZ, PETER
 APPLICANT: HANSEN, PETER KAMP
 TITLE OF INVENTION: PROTEASE VARIANTS AND COMPOSITIONS
 FILE REFERENCE: 4946,200-US
 CURRENT APPLICATION NUMBER: US/08/963,851
 CURRENT FILING DATE: 1997-11-04
 NUMBER OF SEQ ID NOS: 35
 SOFTWARE: FastSeq for Windows Version 3.0
 SEQ ID NO 15
 LENGTH: 68
 TYPE: PRT
 ORGANISM: Staphylococcus epidermis
 US-08-963-851-15

Query Match 100.0%; Score 23; DB 4; Length 68;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DYDA 4
 Db 12 DYDA 15

RESULT 29
 US-09-134-001C-3420
 Sequence 3420, Application US/09134001C
 Patent No. 6380370
 GENERAL INFORMATION:
 APPLICANT: Lynn Doucette-Stamm et al
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
 TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
 FILE REFERENCE: CTC-007
 CURRENT APPLICATION NUMBER: US/09/134,001C
 CURRENT FILING DATE: 1998-08-13
 PRIOR APPLICATION NUMBER: US 60/064,964
 PRIOR FILING DATE: 1997-11-08
 PRIOR APPLICATION NUMBER: US 60/055,779
 PRIOR FILING DATE: 1997-08-14
 NUMBER OF SEQ ID NOS: 5674
 SEQ ID NO 3420
 LENGTH: 79
 TYPE: PRT
 ORGANISM: Staphylococcus epidermidis
 US-09-134-001C-3420

Query Match 100.0%; Score 23; DB 4; Length 79;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 DYDA 4
 Db 47 DYDA 50
 RESULT 30
 US-08-096-741-7
 Sequence 7, Application US/08096741
 Patent No. 5453374
 GENERAL INFORMATION:
 APPLICANT: Furuva, Kaoru
 APPLICANT: Matsuda, Akio
 TITLE OF INVENTION: A Transformant Capable of Producing
 TITLE OF INVENTION: D-Amino Acid Oxidase
 NUMBER OF SEQUENCES: 17
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Pennie & Edmonds
 STREET: 1155 Avenue of the Americas
 CITY: New York
 STATE: New York
 COUNTRY: U.S.A.
 ZIP: 10036-2711
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION NUMBER: US/08/096,741
 FILING DATE: 23-JUL-1993
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Miller, Charles E.
 REGISTRATION NUMBER: 24,576
 REFERENCE/DOCKET NUMBER: 408-089
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 212 790-9090
 TELEFAX: 212869-9741/8864
 TELEX: 66141 PENNIE
 INFORMATION FOR SEQ ID NO: 7:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 83 amino acids
 TYPE: amino acid
 TOPOLOGY: unknown
 MOLECULE TYPE: protein
 US-08-096-741-7

Query Match 100.0%; Score 23; DB 1; Length 83;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DYDA 4
 Db 62 DYDA 65

RESULT 31
 US-09-109-266-4
 Sequence 4, Application US/09109266
 Patent No. 5970111
 GENERAL INFORMATION:
 APPLICANT: Weng, Zude
 APPLICANT: Rafalski, Antoni J.
 APPLICANT: Klein, Theodore M.
 APPLICANT: Miao, Guo-Hua
 APPLICANT: Sakai, Hajime
 TITLE OF INVENTION: Plant Cell Proliferation-Associated Proteins
 FILE REFERENCE: BB-1275-P1
 CURRENT APPLICATION NUMBER: US/09/109,266
 CURRENT FILING DATE: 1998-11-20
 NUMBER OF SEQ ID NOS: 22

; SOFTWARE: Microsoft Word Version 7.0A
; SEQ ID NO 4
; LENGTH: 103
; TYPE: PRT
; ORGANISM: Zea mays
US-09-109-266-4

Query Match 100.0%; Score 23; DB 2; Length 103;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
||||
DB 42 DYDA 45

RESULT 32
PCT-US95-07372-10
; Sequence 10, Application PC/TUS9507372
; GENERAL INFORMATION:
; APPLICANT: Oklahoma Medical Research Foundation
; TITLE OF INVENTION: Calcium Binding Recombinant
; TITLE OF INVENTION: Antibody Against Protein C
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 2800 One Atlantic Center
; STREET: 1201 West Peachtree Street
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30309-3450
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER:
; APPLICATION NUMBER: PCT/US95/07372
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: OMRFI06CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 873-8794
; TELEFAX: (404) 873-8795
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 139 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: Internal
; ORIGINAL SOURCE:
; ORGANISM: HPC-4 Heavy Chain Variable Region (VH Gamma)
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 20..139
; OTHER INFORMATION: /note= "Gln at position 20 starts
; OTHER INFORMATION: mature peptide."
PCT-US95-07372-10

Query Match 100.0%; Score 23; DB 5; Length 139;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
||||

DB 122 DYDA 125

RESULT 33
US-09-228-246-5
; Sequence 5, Application US/09228246
; Patent No. 6245510
; GENERAL INFORMATION:
; APPLICANT: Staskawicz, B. S. et al.
; TITLE OF INVENTION: PRF Protein and Nucleic Acid Sequences: Compositions
; TITLE OF INVENTION: and Methods for Plant Pathogen Resistance
; FILE REFERENCE: 51700
; CURRENT APPLICATION NUMBER: US/09/228,246
; CURRENT FILING DATE: 1999-01-11
; EARLIER APPLICATION NUMBER: 08/680,327
; EARLIER FILING DATE: 1996-07-11
; EARLIER APPLICATION NUMBER: 08/310,912
; EARLIER FILING DATE: 1994-09-22
; EARLIER APPLICATION NUMBER: 08/227,360
; EARLIER FILING DATE: 1994-04-13
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 143
; TYPE: PRT
; ORGANISM: Lycopersicon esculentum
US-09-228-246-5

Query Match 100.0%; Score 23; DB 4; Length 143;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
||||
DB 131 DYDA 134

RESULT 34
US-09-518-046-20
; Sequence 20, Application US/09518046
; Patent No. 6294663
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; APPLICANT: Underwood, Lowell J.
; TITLE OF INVENTION: Transmembrane Serine Protease Overexpressed
; TITLE OF INVENTION: in Ovarian Carcinoma and Uses Thereof
; FILE REFERENCE: D6192CIP
; CURRENT APPLICATION NUMBER: US/09/518,046
; CURRENT FILING DATE: 2000-03-02
; EARLIER APPLICATION NUMBER: 09/261,416
; EARLIER FILING DATE: 1999-03-03
; NUMBER OF SEQ ID NOS: 153
; SEQ ID NO 20
; LENGTH: 149
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; OTHER INFORMATION: protease domain of protease M (ProM)
US-09-518-046-20

Query Match 100.0%; Score 23; DB 4; Length 149;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
||||
DB 43 DYDA 46

RESULT 35
US-09-261-416-5
; Sequence 5, Application US/09261416A


```

; Patent No. 6291663
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; APPLICANT: Underwood, Lowell J.
; TITLE OF INVENTION: TADG-12: A No. 6291663el Transmembrane Serine Protease
; TITLE OF INVENTION: Overexpressed in Ovarian Carcinoma
; FILE REFERENCE: D6192
; CURRENT APPLICATION NUMBER: US/09/261,416A
; CURRENT FILING DATE: 1999-03-03
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 5
; LENGTH: 154
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Serine protease catalytic domain of protease M
; OTHER INFORMATION: (Prom) homologous to similar domain in TADG-12.
US-09-261-416-5

Query Match      100.0%; Score 23; DB 4; Length 154;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
Db 51 DYDA 54

RESULT 36
US-08-858-207A-509
; Sequence 509, Application US/08858207A
; Patent No. 6348328
; GENERAL INFORMATION:
; APPLICANT: Black, Michael
; APPLICANT: Hodgson, John
; APPLICANT: Knowles, David
; APPLICANT: Nicholas, Richard
; APPLICANT: Stodola, Robert
; TITLE OF INVENTION: No. 6348328el Compounds
; NUMBER OF SEQUENCES: 552
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Smithkline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/858,207A
; FILING DATE: 09-MAY-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/017670
; FILING DATE: 14-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Gimmi, Edward R
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: P50475
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 509:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 159 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear

```

```

; MOLECULE TYPE: NO. 6348328e
US-08-858-207A-509

Query Match      100.0%; Score 23; DB 4; Length 159;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
Db 12 DYDA 15

RESULT 37
US-08-560-098A-53
; Sequence 53, Application US/08560098A
; Patent No. 5976841
; GENERAL INFORMATION:
; APPLICANT: WRENDET, Stephan
; APPLICANT: HEINZEL-WIELAND, Regina
; APPLICANT: STEFFENS, Gerd Josef
; TITLE OF INVENTION: Proteins having Fibrinolytic and
; TITLE OF INVENTION: Coagulation-inhibiting Properties
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
; STREET: 1200 G Street, N.W., Suite 700
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/560,098A
; FILING DATE: 17-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: P 44 40 892.7
; FILING DATE: 17-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: EVANS, Joseph D.
; REGISTRATION NUMBER: 26,269
; REFERENCE/DOCKET NUMBER: 148/42448
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-8800
; TELEFAX: (202) 628-8844
; INFORMATION FOR SEQ ID NO: 53:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 171 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-560-098A-53

Query Match      100.0%; Score 23; DB 2; Length 171;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
Db 64 DYDA 67

RESULT 38
US-08-361-395-1
; Sequence 1, Application US/08361395
; Patent No. 5733768
; GENERAL INFORMATION:
; APPLICANT: Dixon, Eric P.
; APPLICANT: Johnstone, Edward M.

```

; APPLICANT: Little, Sheila P.
; APPLICANT: NO. 5733768ris, Franklin H.
; TITLE OF INVENTION: PROTEASE AND RELATED DNA COMPOUNDS
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eli Lilly and Company
; STREET: Lilly Corporate Center/Patent Division
; CITY: Indianapolis
; STATE: Indiana
; COUNTRY: United States
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/361,395
; FILING DATE: 22-DEC-1994
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Bialock, Donna K.
; REGISTRATION NUMBER: 38,082
; REFERENCE/DOCKET NUMBER: X8350A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 317-277-1090
; TELEFAX: 317-276-3861
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 244 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-361-395-1

Query Match 100.0%; Score 23; DB 1; Length 244;
Best Local Similarity 100.0%; Pred. No. 4.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
Db 97 DYDA 100

RESULT 39
US-08-637-759B-92
; Sequence 92, Application US/08637759B
; Patent No. 5876931
; GENERAL INFORMATION:
; APPLICANT: David William Holden
; TITLE OF INVENTION: Identification of Genes
; NUMBER OF SEQUENCES: 501
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 2800 One Atlantic Center
; STREET: 1201 West Peachtree Street
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30309-3450
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/637,759B
; FILING DATE: 03-MAY-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB95/02875
; FILING DATE: 11-DEC-1995

; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: RPMS 101
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 873-8794
; TELEFAX: (404) 873-8795
; INFORMATION FOR SEQ ID NO: 92:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 257 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; US-08-637-759B-92

Query Match 100.0%; Score 23; DB 2; Length 257;
Best Local Similarity 100.0%; Pred. No. 4.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
Db 60 DYDA 63

RESULT 40
US-08-871-355A-92
; Sequence 92, Application US/08871355A
; Patent No. 6015669
; GENERAL INFORMATION:
; APPLICANT: David William Holden
; TITLE OF INVENTION: Identification of Genes
; NUMBER OF SEQUENCES: 501
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 2800 One Atlantic Center
; STREET: 1201 West Peachtree Street
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30309-3450
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/871,355A
; FILING DATE: 09-JUN-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB95/02875
; FILING DATE: 11-DEC-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: RPMS 101 CON
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 873-8794
; TELEFAX: (404) 873-8795
; INFORMATION FOR SEQ ID NO: 92:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 257 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; US-08-871-355A-92

Query Match 100.0%; Score 23; DB 3; Length 257;
Best Local Similarity 100.0%; Pred. No. 4.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DYDA 4
||||
Db 60 DYDA 63

RESULT 41
US-09-201-945-92
; Sequence 92, Application US/09201945
; Patent No. 6342215
; GENERAL INFORMATION:
; APPLICANT: David William Holden
; TITLE OF INVENTION: Identification of Genes
; NUMBER OF SEQUENCES: 501
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 2800 One Atlantic Center
; STREET: 1201 West Peachtree Street
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30309-3450
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/201,945
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/637,759
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.,
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: RPS 101
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 873-8794
; TELEFAX: (404) 873-8795
; INFORMATION FOR SEQ ID NO: 92:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 257 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
US-09-201-945-92

Query Match 100.0%; Score 23; DB 4; Length 257;
Best Local Similarity 100.0%; Pred. No. 4.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DYDA 4
||||
Db 60 DYDA 63

RESULT 42
US-08-312-949-4
; Sequence 4, Application US/08312949
; Patent No. 6027734
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: Wu, Hong-Yin
; TITLE OF INVENTION: MUCOSAL ADMINISTRATION OF
; PNEUMOCOCCAL ANTIGENS

; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/312,949
; FILING DATE: 30-SEP-1994
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2049
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-0712
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 288 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-312-949-4

Query Match 100.0%; Score 23; DB 3; Length 288;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DYDA 4
||||
Db 14 DYDA 17

RESULT 43
US-08-446-201-4
; Sequence 4, Application US/08446201B
; Patent No. 6042838
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: Wu, Hong-Yin
; TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS FOR MUCOSAL ADMINISTRATION OF
; PNEUMOCOCCAL SURFACE PROTEIN A (PSPA)
; FILE REFERENCE: 454312-2018
; CURRENT APPLICATION NUMBER: US/08/446,201B
; CURRENT FILING DATE: 1995-05-19
; EARLIER APPLICATION NUMBER: 08/312,949
; EARLIER FILING DATE: 1994-09-30
; EARLIER APPLICATION NUMBER: 08/246,636
; EARLIER FILING DATE: 1994-05-20
; EARLIER APPLICATION NUMBER: 08/048,896
; EARLIER FILING DATE: 1993-04-20
; EARLIER APPLICATION NUMBER: 07/835,698
; EARLIER FILING DATE: 1992-02-12
; EARLIER APPLICATION NUMBER: 07/656,773
; EARLIER FILING DATE: 1991-02-15
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 288
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-08-446-201-4

Query Match 100.0%; Score 23; DB 3; Length 288;

```
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
Db 14 DYDA 17

RESULT 44
US-08-072-070-4
; Sequence 4, Application US/08072070
; Patent No. 5476929
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: Yother, Janet L
; APPLICANT: McDaniel, Larry S
; TITLE OF INVENTION: STRUCTURAL GENE OF PNEUMOCOCCAL
; TITLE OF INVENTION: PROTEIN
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Shoemaker and Mattare, Ltd
; STREET: Suite 1203, 2001 Jefferson Davis Highway
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202-0286
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/072,070
; FILING DATE: 19930603
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/835,698
; FILING DATE: 12-FEB-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/656,773
; FILING DATE: 15-FEB-1991
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 415-0810
; TELEX: LUKPAT WASHINGTON
; INFORMATION FOR SEQ ID NO: 4:
; LENGTH: 289 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; FRAGMENT TYPE: N-terminal
US-08-072-070-4

Query Match 100.0%; Score 23; DB 1; Length 289;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
Db 14 DYDA 17

RESULT 45
US-08-469-434-4
; Sequence 4, Application US/08469434
; Patent No. 5753463
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: Yother, Janet L
; APPLICANT: McDaniel, Larry S
; TITLE OF INVENTION: STRUCTURAL GENE OF PNEUMOCOCCAL
; TITLE OF INVENTION: PROTEIN
```

```
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Shoemaker and Mattare, Ltd
STREET: Suite 1203, 2001 Jefferson Davis Highway
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202-0286
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/469,434
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/072,065
FILING DATE: 03 JUNE 1993
APPLICATION NUMBER: US/07/835,698
FILING DATE: 12-FEB-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/656,773
FILING DATE: 15-FEB-1991
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 415-0810
TELEFAX: (703) 521-0378
TELEX: LUKPAT WASHINGTON
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 289 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
FRAGMENT TYPE: N-terminal
US-08-469-434-4

Query Match 100.0%; Score 23; DB 1; Length 289;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
Db 14 DYDA 17

Search completed: February 6, 2003, 11:24:03
Job time : 11.3333 secs
```

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 6, 2003, 11:22:09 : Search time 5.83333 Seconds
(without alignments)
15.202 Million cell updates/sec

Title: PAT943-1
Perfect score: 23
Sequence: 1 dyda 4

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 129505 seqs, 22169297 residues
Total number of hits satisfying chosen parameters: 126

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 100%
Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	23	100.0	38	10	US-09-879-957-102
2	23	100.0	38	10	US-09-864-761-4554
3	23	100.0	39	10	US-09-879-957-104
4	23	100.0	40	10	US-09-879-957-96
5	23	100.0	68	10	US-09-948-080-15
6	23	100.0	69	10	US-09-864-761-4134
7	23	100.0	97	9	US-09-738-626-5597
8	23	100.0	143	10	US-09-881-823-16
9	23	100.0	144	10	US-09-796-294-1
10	23	100.0	155	9	US-09-833-790-252
11	23	100.0	165	9	US-09-738-626-4707
12	23	100.0	165	10	US-09-925-301-998
13	23	100.0	184	9	US-09-738-626-6546
14	23	100.0	201	10	US-09-764-853-827
15	23	100.0	221	9	US-09-712-363-211
16	23	100.0	244	10	US-09-796-294-11
17	23	100.0	252	9	US-09-738-626-4443
18	23	100.0	252	10	US-09-987-446-2
19	23	100.0	267	10	US-09-815-242-5328

20	23	100.0	301	9	US-09-971-536-51
21	23	100.0	308	10	US-09-815-242-5306
22	23	100.0	308	10	US-09-815-242-12632
23	23	100.0	318	10	US-09-815-242-13749
24	23	100.0	323	10	US-09-815-242-12514
25	23	100.0	323	10	US-09-815-242-12857
26	23	100.0	326	9	US-10-033-297-75
27	23	100.0	326	9	US-10-033-297-135
28	23	100.0	326	10	US-09-777-4308-28
29	23	100.0	330	9	US-09-738-626-5624
30	23	100.0	339	9	US-09-738-626-6267
31	23	100.0	341	10	US-09-815-242-10772
32	23	100.0	344	10	US-09-925-300-1154
33	23	100.0	365	10	US-09-741-669-353
34	23	100.0	394	10	US-09-925-301-1388
35	23	100.0	399	10	US-09-764-870-504
36	23	100.0	415	9	US-09-738-626-3589
37	23	100.0	417	10	US-09-815-242-10254
38	23	100.0	419	10	US-09-815-242-13792
39	23	100.0	432	10	US-09-815-242-5096
40	23	100.0	437	10	US-09-765-272-70
41	23	100.0	438	9	US-09-738-626-6703
42	23	100.0	446	12	US-10-081-859-2
43	23	100.0	455	9	US-09-738-626-6527
44	23	100.0	472	10	US-09-815-242-13727
45	23	100.0	480	10	US-09-815-242-13367

ALIGNMENTS

RESULT 1
US-09-879-957-102
; Sequence 102, Application US/09879957
; Patent No. US20020034755A1
; GENERAL INFORMATION:
; APPLICANT: SPARKS, Andrew B.
; HOFFMAN, No. US20020034755A1h
; KAY, Brian K.
; FOWLES, Dana M.
; MCCONNELL, Stephen J.
; TITLE OF INVENTION: POLYPEPTIDES HAVING A FUNCTIONAL
; DOMAIN OF INTEREST AND METHODS OF IDENTIFYING AND
; USING SAME
; NUMBER OF SEQUENCES: 227
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/879,957
; FILING DATE: 13-Jun-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/630,915
; FILING DATE: 03-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Misrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 1101-174
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 102:

SEQUENCE CHARACTERISTICS:
LENGTH: 38 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: unknown
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 102:
US-09-879-957-102

Query Match 100.0%; Score 23; DB 10; Length 38;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DYDA 4
Db 5 DYDA 8

RESULT 2
US-09-864-761-45554
; Sequence 45554, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263,6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 45554
; LENGTH: 38
; TYPE: PRT

ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO ALI57392.2
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.96
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.83
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.74
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.88
OTHER INFORMATION: EST_HUMAN HIT: BF574717.1, EVALUE 1.00e-14
OTHER INFORMATION: SWISSPROT HIT: P08442, EVALUE 7.40e-01
US-09-864-761-45554

Query Match 100.0%; Score 23; DB 10; Length 38;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DYDA 4
Db 19 DYDA 22

RESULT 3
US-09-879-957-104
; Sequence 104, Application US/09879957
; Patent No. US20020034755A1
; GENERAL INFORMATION:
; APPLICANT: SPARKS, Andrew B.
; HOFFMAN, No. US20020034755A1
; KAY, Brian K.
; FOWLKES, Dana M.
; MCCONNELL, Stephen J.
; TITLE OF INVENTION: POLYPEPTIDES HAVING A FUNCTIONAL
; DOMAIN OF INTEREST AND METHODS OF IDENTIFYING AND
; USING SAME
; NUMBER OF SEQUENCES: 227
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/879,957
; FILING DATE: 13-Jun-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/630,915
; FILING DATE: 03-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Misrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 1101-174
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 104:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 39 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 104:
US-09-879-957-104

Query Match 100.0%; Score 23; DB 10; Length 39;

Best Local Similarity 100.0%; Pred. No. 41;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DYDA 4
Db 5 DYDA 8

RESULT 4

US-09-879-957-96
; Sequence 96, Application US/09879957
; Patent No. US20020034755A1
; GENERAL INFORMATION:
; APPLICANT: SPARKS, Andrew B.
; HOFFMAN, No. US20020034755A1h
; KAY, Brian K.
; FOWLKES, Dana M.
; MCCONNELL, Stephen J.

TITLE OF INVENTION: POLYPEPTIDES HAVING A FUNCTIONAL
DOMAIN OF INTEREST AND METHODS OF IDENTIFYING AND
USING SAME

NUMBER OF SEQUENCES: 227

CORRESPONDENCE ADDRESS:

ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/879,957

FILING DATE: 13-Jun-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/630,915

FILING DATE: 03-APR-1996

ATTORNEY/AGENT INFORMATION:

NAME: Mistock, S. Leslie

REGISTRATION NUMBER: 18,872

REFERENCE/DOCKET NUMBER: 1101-174

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 790-9090

TELEFAX: (212) 869-8864/9741

INFORMATION FOR SEQ ID NO: 96:

SEQUENCE CHARACTERISTICS:

LENGTH: 40 amino acids

TYPE: amino acid

STRANDEDNESS: <Unknown>

TOPOLOGY: unknown

MOLECULE TYPE: peptide

SEQUENCE DESCRIPTION: SEQ ID NO: 96:

US-09-879-957-96

Query Match 100.0%; Score 23; DB 10; Length 40;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DYDA 4
Db 5 DYDA 8

RESULT 5

US-09-948-080-15

; Sequence 15, Application US/09948080

; Patent No. US20020102702A1

; GENERAL INFORMATION:

; APPLICANT: VAN DER OSTEN, CLAUS
; APPLICANT: HALKIER, TORDEN
; APPLICANT: ANDERSEN, CARSTEN
; APPLICANT: BAUDITZ, PETER
; APPLICANT: HANSEN, PETER KAMP
; TITLE OF INVENTION: PROTEASE VARIANTS AND COMPOSITIONS
; FILE REFERENCE: 4946,200-US
; CURRENT APPLICATION NUMBER: US/09/948,080
; CURRENT FILING DATE: 2001-09-06
; PRIOR APPLICATION NUMBER: US/08/963,851
; PRIOR FILING DATE: 1997-11-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 15
; LENGTH: 68
; TYPE: PRT
; ORGANISM: Staphylococcus epidermis

US-09-948-080-15

Query Match 100.0%; Score 23; DB 10; Length 68;

Best Local Similarity 100.0%; Pred. No. 70;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DYDA 4

Db 12 DYDA 15

RESULT 6

US-09-864-761-41344

; Sequence 41344, Application US/09864761

; Patent No. US20020048763A1

; GENERAL INFORMATION:

; APPLICANT: Penn, Sharron G.

; APPLICANT: Rank, David R.

; APPLICANT: Hanzel, David K.

; APPLICANT: Chen, Wensheng

; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL

; FILE REFERENCE: Aemica-X-1

; CURRENT APPLICATION NUMBER: US/09/864,761

; CURRENT FILING DATE: 2001-05-23

; PRIOR APPLICATION NUMBER: US 60/180,312

; PRIOR FILING DATE: 2000-02-04

; PRIOR APPLICATION NUMBER: US 60/207,456

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: US 09/632,366

; PRIOR FILING DATE: 2000-08-03

; PRIOR APPLICATION NUMBER: GB 24263.6

; PRIOR FILING DATE: 2000-10-04

; PRIOR APPLICATION NUMBER: US 60/236,359

; PRIOR FILING DATE: 2000-09-27

; PRIOR APPLICATION NUMBER: PCT/US01/00666

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00667

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00664

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00669

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00665

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00668

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00663

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00662

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00661

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00670

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: US 60/234,687

; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 41344
; LENGTH: 69
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL121895.18
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.1
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.1
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.2
; OTHER INFORMATION: SWISSPROT HIT: P48193, EVALUE 2.00e-24
; OTHER INFORMATION: EST_HUMAN HIT: BF347885.1, EVALUE 7.00e-34
US-09-864-761-41344

Query Match 100.0%; Score 23; DB 10; Length 69;
Best Local Similarity 100.0%; Pred. No. 71;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
||||
Db 34 DYDA 37

RESULT 7
US-09-738-626-5597
; Sequence 5597, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/37484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 5597
; LENGTH: 97
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-5597

Query Match 100.0%; Score 23; DB 9; Length 97;
Best Local Similarity 100.0%; Pred. No. 99;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
||||
Db 4 DYDA 7

RESULT 8
US-09-881-823-16
; Sequence 16, Application US/09881823
; Patent No. US20020068066A1
; GENERAL INFORMATION:
; APPLICANT: SHI, WENYUAN
; APPLICANT: ANDERSON, MAXWELL
; APPLICANT: MORRISON, SHERIE
; APPLICANT: TRINH, RYAN
; APPLICANT: WIMS, LETITIA
; APPLICANT: CHEN, LI
; TITLE OF INVENTION: Method for the Treatment and Prevention of Dental Caries
; FILE REFERENCE: 22851-032
; CURRENT APPLICATION NUMBER: US/09/881,823
; CURRENT FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 07/378,577
; PRIOR FILING DATE: 1999-08-20
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 16
; LENGTH: 143
; TYPE: PRT
; ORGANISM: Murine
US-09-881-823-16

Query Match 100.0%; Score 23; DB 10; Length 143;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
||||
Db 123 DYDA 126

RESULT 9
US-09-796-294-1
; Sequence 1, Application US/09796294
; Patent No. US20020037581A1
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; APPLICANT: Underwood, Lowell J.
; TITLE OF INVENTION: Extracellular Serine Protease
; FILE REFERENCE: D6020CIP3
; CURRENT APPLICATION NUMBER: US/09/796,294
; CURRENT FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 09/618,259
; PRIOR FILING DATE: 2000-07-18
; NUMBER OF SEQ ID NOS: 72
; SEQ ID NO 1
; LENGTH: 144
; TYPE: PRT
; ORGANISM: unknown
; FEATURE:
; NAME/KEY: DOMAIN
; OTHER INFORMATION: Amino acid sequence of Protease m (Prom) catalytic
; OTHER INFORMATION: domain
US-09-796-294-1

Query Match 100.0%; Score 23; DB 10; Length 144;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
||||
Db 42 DYDA 45

RESULT 10
US-09-833-790-252
; Sequence 252, Application US/09833790
; Patent No. US20020068288A1

; GENERAL INFORMATION:
; APPLICANT: Lodes, Michael J.
; APPLICANT: Wang, Tongtong
; APPLICANT: Secrist, Heather
; APPLICANT: Mohamath, Raedoh
; APPLICANT: Indrias, Carol Y.
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.512
; CURRENT APPLICATION NUMBER: US/09/833,790
; CURRENT FILING DATE: 2001-04-11
; NUMBER OF SEQ ID NOS: 440
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 252
; LENGTH: 156
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-833-790-252

Query Match 100.0%; Score 23; DB 10; Length 156;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DYDA 4
||||
Db 141 DYDA 144

RESULT 11
US-09-738-626-4707
; Sequence 4707, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 4707
; LENGTH: 165
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-4707

Query Match 100.0%; Score 23; DB 9; Length 165;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DYDA 4
||||
Db 156 DYDA 159

RESULT 12
US-09-925-301-998

; Sequence 998, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 998
; LENGTH: 165
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-301-998

Query Match 100.0%; Score 23; DB 10; Length 165;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DYDA 4
||||
Db 152 DYDA 155

RESULT 13
US-09-738-626-6546
; Sequence 6546, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 6546
; LENGTH: 184
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-6546

Query Match 100.0%; Score 23; DB 9; Length 184;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DYDA 4
||||
Db 67 DYDA 70

RESULT 14
US-09-764-853-927

```
; Sequence 827, Application US/09764853
; Patent No. US20020090672A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PJ206
; CURRENT APPLICATION NUMBER: US/09/764,853
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 939
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 827
; LENGTH: 201
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (182)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-853-827

Query Match
Best Local Similarity 100.0%; Score 23; DB 10; Length 201;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
   ||||
Db 189 DYDA 192

RESULT 15
US-09-712-363-211
; Sequence 211, Application US/09712363
; Patent No. US20020164588A1
; GENERAL INFORMATION:
; APPLICANT: Eisenberg, David
; APPLICANT: Rotstein, Sergio H.
; APPLICANT: Marcotte, Edward M.
; TITLE OF INVENTION: DETERMINING THE FUNCTIONS AND
; INTERACTIONS OF PROTEINS BY COMPARATIVE ANALYSIS
; FILE REFERENCE: 07419-032001
; CURRENT APPLICATION NUMBER: US/09/712,363
; CURRENT FILING DATE: 2000-11-13
; PRIOR APPLICATION NUMBER: PCT/US00/02246
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/179,531
; PRIOR FILING DATE: 1999-01-29
; PRIOR APPLICATION NUMBER: 60/117,844
; PRIOR FILING DATE: 1999-01-29
; PRIOR APPLICATION NUMBER: 60/118,206,
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: 60/126,593
; PRIOR FILING DATE: 1999-03-26
; PRIOR APPLICATION NUMBER: 60/134,093
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: 60/134,092
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: 60/165,124
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/165,086
; PRIOR FILING DATE: 1999-11-12
; NUMBER OF SEQ ID NOS: 292
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 211
; LENGTH: 221
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-09-712-363-211

Query Match
Best Local Similarity 100.0%; Score 23; DB 9; Length 221;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 DYDA 4
   ||||
Db 191 DYDA 194

RESULT 16
US-09-796-294-11
; Sequence 11, Application US/09796294
; Patent No. US20020037581A1
; GENERAL INFORMATION:
; APPLICANT: Underwood, Timothy J.
; APPLICANT: O'Brien, Lowell J.
; TITLE OF INVENTION: Extracellular Serine Protease
; FILE REFERENCE: D6020CIP3
; CURRENT APPLICATION NUMBER: US/09/796,294
; CURRENT FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 09/618,259
; PRIOR FILING DATE: 2000-07-18
; NUMBER OF SEQ ID NOS: 72
; SEQ ID NO 11
; LENGTH: 244
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Amino acid sequence of human protease m (hProm);
; OTHER INFORMATION: accession no. U62801
US-09-796-294-11
```

```
Query Match
Best Local Similarity 100.0%; Score 23; DB 10; Length 244;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 DYDA 4
   ||||
Db 97 DYDA 100

RESULT 17
US-09-738-626-4443
; Sequence 443, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 443
; LENGTH: 252
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-4443
```

```
Query Match
Best Local Similarity 100.0%; Score 23; DB 9; Length 252;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 1 DYDA 4
||||
Db 178 DYDA 181

RESULT 18

US-09-987-446-2
; Sequence 2, Application US/09987446
; Patent No. US20020115162A1
; GENERAL INFORMATION:
; APPLICANT: FARWICK, Mike, et al.
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES CODING FOR THE CYSQ GENE
; FILE REFERENCE: 032301 WD 245
; CURRENT APPLICATION NUMBER: US/09/987,446
; CURRENT FILING DATE: 2001-11-14
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 252
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-987-446-2

Query Match 100.0%; Score 23; DB 10; Length 252;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
||||
Db 178 DYDA 181

RESULT 19

US-09-815-242-5328
; Sequence 5328, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5328
; LENGTH: 267
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-5328

Query Match 100.0%; Score 23; DB 10; Length 267;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
||||
Db 118 DYDA 121

RESULT 20

US-09-971-536-51
; Sequence 51, Application US/09971536
; Patent No. US20020159976A1
; GENERAL INFORMATION:
; APPLICANT: Glenn, Matthew
; APPLICANT: Havukkala, Ilkka
; APPLICANT: Bliksberg, Leonard
; APPLICANT: Lubbers, Mark
; APPLICANT: Dekker, James
; APPLICANT: Christensson, Anna
; APPLICANT: Holland, Ross
; APPLICANT: O'Toole, Paul
; APPLICANT: Reid, Julian
; APPLICANT: Coolbear, Timothy
; TITLE OF INVENTION: Lactobacillus rhamnosus Polynucleotides, Polypeptides and
; FILE REFERENCE: 1043C2
; CURRENT APPLICATION NUMBER: US/09/971,536
; CURRENT FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: U.S. No. US20020159976A1 09/634,238
; PRIOR FILING DATE: 2000-08-08
; PRIOR APPLICATION NUMBER: U.S. No. US20020159976A1 09/724,623
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: PCT/NZ01/00160
; PRIOR FILING DATE: 2001-08-08
; NUMBER OF SEQ ID NOS: 83
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 51
; LENGTH: 301
; TYPE: PRT
; ORGANISM: Lactobacillus rhamnosus
US-09-971-536-51

Query Match 100.0%; Score 23; DB 9; Length 301;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
||||
Db 64 DYDA 67

RESULT 21

US-09-815-242-5306
; Sequence 5306, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848

;
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5306
; LENGTH: 308
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-5306

Query Match 100.0%; Score 23; DB 10; Length 308;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DYDA 4

Db 161 DYDA 164

RESULT 22
US-09-815-242-12632
; Sequence 12632, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR APPLICATION NUMBER: 60/269,308
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12632
; LENGTH: 308
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-12632

Query Match 100.0%; Score 23; DB 10; Length 308;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DYDA 4

Db 161 DYDA 164
RESULT 23
US-09-815-242-13749
; Sequence 13749, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR APPLICATION NUMBER: 60/269,308
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13749
; LENGTH: 318
; TYPE: PRT
; ORGANISM: Salmonella typhi
; NAME/KEY: VARIANT
; LOCATION: (1)...(318)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-815-242-13749

Query Match 100.0%; Score 23; DB 10; Length 318;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DYDA 4

Db 168 DYDA 171

RESULT 24
US-09-815-242-12514
; Sequence 12514, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A

; CURRENT APPLICATION NUMBER: US/09/815,242
 ; CURRENT FILING DATE: 2001-03-21
 ; PRIOR APPLICATION NUMBER: 60/191,078
 ; PRIOR FILING DATE: 2000-03-21
 ; PRIOR APPLICATION NUMBER: 60/206,848
 ; PRIOR FILING DATE: 2000-05-23
 ; PRIOR APPLICATION NUMBER: 60/207,727
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: 60/242,578
 ; PRIOR FILING DATE: 2000-10-23
 ; PRIOR APPLICATION NUMBER: 60/253,625
 ; PRIOR FILING DATE: 2000-11-27
 ; PRIOR APPLICATION NUMBER: 60/257,931
 ; PRIOR FILING DATE: 2000-12-22
 ; PRIOR APPLICATION NUMBER: 60/269,308
 ; PRIOR FILING DATE: 2001-02-16
 ; NUMBER OF SEQ ID NOS: 14110
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 12514
 ; LENGTH: 323
 ; TYPE: PRT
 ; ORGANISM: Staphylococcus aureus
 US-09-815-242-12514

Query Match 100.0%; Score 23; DB 10; Length 323;
 Best Local Similarity 100.0%; Pred. No. 3.2e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
 ||||
 DB 167 DYDA 170

RESULT 25
 US-09-815-242-12857
 ; Sequence 12857, Application US/09815242
 ; Patent No. US20020061569A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Haselbeck, Robert
 ; APPLICANT: Ohlsen, Karl L.
 ; APPLICANT: Zyskind, Judith W.
 ; APPLICANT: Wall, Daniel
 ; APPLICANT: Trawick, John D.
 ; APPLICANT: Carr, Grant J.
 ; APPLICANT: Yamamoto, Robert T.
 ; APPLICANT: Xu, H. Howard
 ; TITLE OF INVENTION: Identification of Essential Genes in
 ; FILE REFERENCE: ELITRA.011A
 ; CURRENT APPLICATION NUMBER: US/09/815,242
 ; CURRENT FILING DATE: 2001-03-21
 ; PRIOR APPLICATION NUMBER: 60/191,078
 ; PRIOR FILING DATE: 2000-03-21
 ; PRIOR APPLICATION NUMBER: 60/206,848
 ; PRIOR FILING DATE: 2000-05-23
 ; PRIOR APPLICATION NUMBER: 60/207,727
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: 60/242,578
 ; PRIOR FILING DATE: 2000-10-23
 ; PRIOR APPLICATION NUMBER: 60/253,625
 ; PRIOR FILING DATE: 2000-11-27
 ; PRIOR APPLICATION NUMBER: 60/257,931
 ; PRIOR FILING DATE: 2000-12-22
 ; PRIOR APPLICATION NUMBER: 60/269,308
 ; PRIOR FILING DATE: 2001-02-16
 ; NUMBER OF SEQ ID NOS: 14110
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 12857
 ; LENGTH: 323
 ; TYPE: PRT
 ; ORGANISM: Staphylococcus aureus
 US-09-815-242-12857

Query Match 100.0%; Score 23; DB 10; Length 323;
 Best Local Similarity 100.0%; Pred. No. 3.2e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DYDA 4
 ||||
 DB 167 DYDA 170

RESULT 26
 US-10-033-297-75
 ; Sequence 75, Application US/10033297
 ; Publication No. US20020187486A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Hall, Jeff G.
 ; Lyamichiev, Victor I.
 ; Mast, Andrea L.
 ; Brow, Mary Ann D.
 ; TITLE OF INVENTION: Detection Of Nucleic Acids By Multiple
 ; Sequential Invasive Cleavages
 ; NUMBER OF SEQUENCES: 163
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Medlen & Carroll, LLP
 ; STREET: 220 Montgomery Street, Suite 2200
 ; CITY: San Francisco
 ; STATE: California
 ; COUNTRY: United States Of America
 ; ZIP: 94104

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/10/033,297
 FILING DATE: 12-No. US20020187486A1-2001
 CLASSIFICATION: <Unknown>
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/09/350,597
 FILING DATE: 09-Jul-1999
 APPLICATION NUMBER: US/08/823,516
 FILING DATE: 24-MAR-1997
 APPLICATION NUMBER: PCT/US97/01072
 FILING DATE: 21-JAN-1997
 APPLICATION NUMBER: US 08/759,038
 FILING DATE: 02-DEC-1996
 APPLICATION NUMBER: US 08/758,314
 FILING DATE: 02-DEC-1996
 APPLICATION NUMBER: US 08/756,386
 FILING DATE: 29-NOV-1996
 APPLICATION NUMBER: US 08/682,853
 FILING DATE: 12-JUL-1996
 APPLICATION NUMBER: US 08/599,491
 FILING DATE: 24-JAN-1996

ATTORNEY/AGENT INFORMATION:
 NAME: Ingolia, Diane E.
 REGISTRATION NUMBER: 40,027
 REFERENCE/DOCKET NUMBER: FORS-02736
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 705-8410
 TELEFAX: (415) 397-8338
 INFORMATION FOR SEQ ID NO: 75:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 326 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 SEQUENCE DESCRIPTION: SEQ ID NO: 75:
 US-10-033-297-75

Query Match 100.0%; Score 23; DB 9; Length 326;
 Best Local Similarity 100.0%; Pred. No. 3.3e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	DYDA	4
Db	173	DYDA	176

```

RESULT 27
US-10-033-297-135
; Sequence 135, Application US/10033297
; Publication No. US20020187486A1
; GENERAL INFORMATION:
; APPLICANT: Hall, Jeff G.
; Dyamichev, Victor I.
; Mast, Andrea L.
; Brow, Mary Ann D.
; TITLE OF INVENTION: Detection Of Nucleic Acids By Multiple
; Sequential Invasive Cleavages
; NUMBER OF SEQUENCES: 163
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medlen & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States Of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/033,297
; FILING DATE: 12-No. US20020187486A1-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/350,597
; FILING DATE: 09-Jul-1999
; APPLICATION NUMBER: US/08/823,516
; FILING DATE: 24-MAR-1997
; APPLICATION NUMBER: PCT/US97/01072
; FILING DATE: 21-JAN-1997
; APPLICATION NUMBER: US 08/759,038
; FILING DATE: 02-DEC-1996
; APPLICATION NUMBER: US 08/758,314
; FILING DATE: 02-DEC-1996
; APPLICATION NUMBER: US 08/756,386
; FILING DATE: 29-NOV-1996
; APPLICATION NUMBER: US 08/682,853
; FILING DATE: 12-JUL-1996
; APPLICATION NUMBER: US 08/599,491
; FILING DATE: 24-JAN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Ingolia, Diane E.
; REGISTRATION NUMBER: 401,027
; REFERENCE/DOCKET NUMBER: FORS-02736
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 135:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 326 amino acids
; TYPE: amino acid
; STRANDEDNESS: No. US20020187486A1 Relevant
; TOPOLOGY: No. US20020187486A1 Relevant
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 135:
US-10-033-297-135

```

```

Db      173 DYDA 176

|||||

RESULT 28
US-09-777-430A-28
; Sequence 28, Application US/09777430A
; Patent No. US20020128465A1
; GENERAL INFORMATION:
; APPLICANT: Lyamichev, Victor
; APPLICANT: Skrzypczynski, Zbigniew
; APPLICANT: Allawi, Hatim T.
; APPLICANT: Wayland, Sarah R.
; APPLICANT: Takova, Tsetska
; APPLICANT: Near, Bruce P.
; TITLE OF INVENTION: Charge Tags and the Separation of Nucleic Acid Molecules
; FILE REFERENCE: FORS-04912
; CURRENT APPLICATION NUMBER: US/09/777,430A
; CURRENT FILING DATE: 2001-02-06
; NUMBER OF SEQ ID NOS: 85
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 28
; LENGTH: 326
; TYPE: PRT
; ORGANISM: Methanococcus jannaschii
US-09-777-430A-28

```

```

RESULT 29
US-09-738-626-5624
; Sequence 5624, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 5624
; LENGTH: 330
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-5624

```

```

Query Match      100.0%; Score 23; DB 9; Length 330;
Best Local Similarity 100.0%; Pred. NO. 3.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 1 DYDA 4
||||
Db 215 DYDA 218

RESULT 30
US-09-738-626-6267
; Sequence 6267, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 6267
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-6267

Query Match 100.0%; Score 23; DB 9; Length 339;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DYDA 4
||||
Db 46 DYDA 49

RESULT 31
US-09-815-242-10772
; Sequence 10772, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23

; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10772
; LENGTH: 341
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-815-242-10772

Query Match 100.0%; Score 23; DB 10; Length 341;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DYDA 4
||||
Db 234 DYDA 237

RESULT 32
US-09-925-300-1154
; Sequence 1154, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925,300
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1154
; LENGTH: 344
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (85)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (88)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (96)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (140)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (314)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-300-1154

Query Match 100.0%; Score 23; DB 10; Length 344;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DYDA 4
||||
Db 228 DYDA 231

RESULT 33
US-09-741-669-353

; Sequence 353, Application US/09741669
; Patent No. US20020022718A1
; GENERAL INFORMATION:
; APPLICANT: Forsyth, R. Allyn
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; TITLE OF INVENTION: Genes identified as required for
; proliferation of E. coli
; FILE REFERENCE: ELITRA.009A
; CURRENT APPLICATION NUMBER: US/09/741,669
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: US 60/173005
; PRIOR FILING DATE: 1999-12-23
; NUMBER OF SEQ ID NOS: 481
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 353
; LENGTH: 365
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-741-669-353

Query Match 100.0%; Score 23; DB 10; Length 365;
Best Local Similarity 100.0%; Pred. No. 3.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
|||||
Db 26 DYDA 29

RESULT 34
US-09-925-301-1388
; Sequence 1388, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1388
; LENGTH: 394
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (3)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-301-1388

Query Match 100.0%; Score 23; DB 10; Length 394;
Best Local Similarity 100.0%; Pred. No. 3.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
|||||
Db 107 DYDA 110

RESULT 35
US-09-764-870-504
; Sequence 504, Application US/09764870
; Patent No. US20020042386A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT214

; CURRENT APPLICATION NUMBER: US/09/764,870
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 646
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 504
; LENGTH: 399
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (22)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (33)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (369)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-870-504

Query Match 100.0%; Score 23; DB 10; Length 399;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
|||||
Db 283 DYDA 286

RESULT 36
US-09-738-626-3589
; Sequence 3589, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIO
; APPLICANT: OCHIALI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 3589
; LENGTH: 415
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-3589

Query Match 100.0%; Score 23; DB 9; Length 415;
Best Local Similarity 100.0%; Pred. No. 4.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
|||||
Db 42 DYDA 45

RESULT 37


```
US-09-815-242-10254
; Sequence 10254, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10254
; LENGTH: 417
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-815-242-10254
```

```
Query Match 100.0%; Score 23; DB 10; Length 417;
Best Local Similarity 100.0%; Pred. No. 4.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 DYDA 4
    ||||
Db 10 DYDA 13
```

```
RESULT 38
US-09-815-242-13792
; Sequence 13792, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
```

```
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13792
; LENGTH: 419
; TYPE: PRT
; ORGANISM: Salmonella typhi
US-09-815-242-13792
```

```
Query Match 100.0%; Score 23; DB 10; Length 419;
Best Local Similarity 100.0%; Pred. No. 4.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 DYDA 4
    ||||
Db 12 DYDA 15
```

```
RESULT 39
US-09-815-242-5096
; Sequence 5096, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5096
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-815-242-5096
```

```
Query Match 100.0%; Score 23; DB 10; Length 432;
Best Local Similarity 100.0%; Pred. No. 4.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 DYDA 4
    ||||
Db 415 DYDA 418
```

```
RESULT 40
US-09-765-272-70
; Sequence 70, Application US/09765272
; Patent No. US20020061545A1
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigen and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/765,272
; FILING DATE: 22-Jan-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/961,083
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 361,373
; REFERENCE/DOCKET NUMBER: PB340P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 70:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 437 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 70:
US-09-765-272-70

Query Match 100.0%; Score 23; DB 10; Length 437;
Best Local Similarity 100.0%; Pred. No. 4.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
Db 288 DYDA 291

RESULT 41
US-09-738-626-6703
; Sequence 6703, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIALI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738, 626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
```

```
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 6703
; LENGTH: 438
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-6703

Query Match 100.0%; Score 23; DB 9; Length 438;
Best Local Similarity 100.0%; Pred. No. 4.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
Db 108 DYDA 111

RESULT 42
US-10-081-859-2
; Sequence 2, Application US/10081859
; Patent No. US20020120122A1
; GENERAL INFORMATION:
; APPLICANT: SUGA, MIKIKO
; APPLICANT: ASAKURA, YOKO
; APPLICANT: SUGIMOTO, MASAKAZU
; APPLICANT: ITO, HISAO
; TITLE OF INVENTION: PHOSPHOSERINE PHOSPHATASE GENE OF CORYNEFORM BACTERIA
; FILE REFERENCE: 202044USO
; CURRENT APPLICATION NUMBER: US/10/081,859
; CURRENT FILING DATE: 2002-02-25
; PRIOR APPLICATION NUMBER: 09/761,716
; PRIOR FILING DATE: 2001-01-18
; PRIOR APPLICATION NUMBER: JP2000-023341
; PRIOR FILING DATE: 2000-01-27
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 446
; TYPE: PRT
; ORGANISM: Brevibacterium flavum
US-10-081-859-2

Query Match 100.0%; Score 23; DB 12; Length 446;
Best Local Similarity 100.0%; Pred. No. 4.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
Db 153 DYDA 156

RESULT 43
US-09-738-626-6527
; Sequence 6527, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIALI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
```

; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 6527
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-6527

Query Match 100.0%; Score 23; DB 9; Length 455;
Best Local Similarity 100.0%; Pred. No. 4.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
Db 226 DYDA 229

RESULT 44
US-09-815-242-13727
; Sequence 13727, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13727
; LENGTH: 472
; TYPE: PRT
; ORGANISM: Salmonella typhi
; NAME/KEY: VARIANT
; LOCATION: (1)...(472)
; OTHER INFORMATION: Xaa - Any Amino Acid
US-09-815-242-13727

Query Match 100.0%; Score 23; DB 10; Length 472;
Best Local Similarity 100.0%; Pred. No. 4.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
Db 231 DYDA 234

RESULT 45
US-09-815-242-13367
; Sequence 13367, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13367
; LENGTH: 480
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-815-242-13367

Query Match 100.0%; Score 23; DB 10; Length 480;
Best Local Similarity 100.0%; Pred. No. 4.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
Db 313 DYDA 316

Search completed: February 6, 2003, 11:39:42
Job time : 7.83333 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 6, 2003, 11:16:59 ; Search time 133.167 Seconds
(without alignments)
19.366 Million cell updates/sec

Title: PAT943-1
Perfect score: 23
Sequence: 1 dyda 4

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 4569144 seqs, 644733110 residues
Total number of hits satisfying chosen parameters: 4667

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 100%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_AA_Main:
1: /cgn2_6/ptodata/1/paa/PCTUS_COMB.pep.*
2: /cgn2_6/ptodata/1/paa/US06_COMB.pep.*
3: /cgn2_6/ptodata/1/paa/US07_COMB.pep.*
4: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
5: /cgn2_6/ptodata/1/paa/US081_COMB.pep.*
6: /cgn2_6/ptodata/1/paa/US082_COMB.pep.*
7: /cgn2_6/ptodata/1/paa/US083_COMB.pep.*
8: /cgn2_6/ptodata/1/paa/US084_COMB.pep.*
9: /cgn2_6/ptodata/1/paa/US085_COMB.pep.*
10: /cgn2_6/ptodata/1/paa/US086_COMB.pep.*
11: /cgn2_6/ptodata/1/paa/US087_COMB.pep.*
12: /cgn2_6/ptodata/1/paa/US088_COMB.pep.*
13: /cgn2_6/ptodata/1/paa/US089_COMB.pep.*
14: /cgn2_6/ptodata/1/paa/US090_COMB.pep.*
15: /cgn2_6/ptodata/1/paa/US091_COMB.pep.*
16: /cgn2_6/ptodata/1/paa/US092_COMB.pep.*
17: /cgn2_6/ptodata/1/paa/US093_COMB.pep.*
18: /cgn2_6/ptodata/1/paa/US094_COMB.pep.*
19: /cgn2_6/ptodata/1/paa/US095_COMB.pep.*
20: /cgn2_6/ptodata/1/paa/US096_COMB.pep.*
21: /cgn2_6/ptodata/1/paa/US097_COMB.pep.*
22: /cgn2_6/ptodata/1/paa/US098_COMB.pep.*
23: /cgn2_6/ptodata/1/paa/US099_COMB.pep.*
24: /cgn2_6/ptodata/1/paa/US100_COMB.pep.*
25: /cgn2_6/ptodata/1/paa/US101_COMB.pep.*
26: /cgn2_6/ptodata/1/paa/US102_COMB.pep.*
27: /cgn2_6/ptodata/1/paa/US60_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	23	100.0	13	14	US-09-057-228-60
2	23	100.0	13	14	US-09-057-416-60
3	23	100.0	13	21	US-09-716-964A-60
4	23	100.0	20	27	US-60-160-209-3652
5	23	100.0	20	27	US-60-169-868-4801
6	23	100.0	21	1	PCT-US00-09866A-10

7	23	100.0	24	27	US-60-160-209-3607
8	23	100.0	24	27	US-60-160-209-2726
9	23	100.0	28	6	US-08-263-921B-36
10	23	100.0	28	7	US-08-346-602A-36
11	23	100.0	28	15	US-09-118-990A-36
12	23	100.0	28	27	US-60-160-209-2539
13	23	100.0	28	27	US-60-169-868-3774
14	23	100.0	31	27	US-60-162-247-4129
15	23	100.0	38	1	PCT-US01-00663-36701
16	23	100.0	38	8	US-08-417-872A-55
17	23	100.0	38	22	US-09-864-761-45554
18	23	100.0	38	22	US-09-879-957-102
19	23	100.0	38	25	US-10-182-993-35614
20	23	100.0	38	26	US-10-203-134-36629
21	23	100.0	38	26	US-10-203-136-36681
22	23	100.0	38	26	US-10-203-137-36701
23	23	100.0	38	26	US-10-203-139-35253
24	23	100.0	38	27	US-60-161-932-1951
25	23	100.0	39	8	US-08-417-872A-57
26	23	100.0	39	22	US-09-879-957-104
27	23	100.0	40	8	US-08-417-872A-49
28	23	100.0	40	22	US-09-879-957-96
29	23	100.0	41	19	US-09-513-996A-9365
30	23	100.0	41	19	US-09-513-996A-68915
31	23	100.0	41	20	US-09-620-394B-5748
32	23	100.0	45	6	US-08-247-491-3
33	23	100.0	45	8	US-08-467-852-3
34	23	100.0	45	8	US-08-470-626-3
35	23	100.0	45	27	US-60-163-062-951
36	23	100.0	46	18	US-09-450-969-3984
37	23	100.0	47	24	US-10-008-524A-154
38	23	100.0	49	10	US-08-651-136A-58
39	23	100.0	49	24	US-10-007-521-58
40	23	100.0	51	27	US-60-177-571-4760
41	23	100.0	52	24	US-10-015-127-10406
42	23	100.0	53	27	US-60-360-039-12882
43	23	100.0	54	24	US-10-029-386-28441
44	23	100.0	58	1	PCT-US01-03800A-1476
45	23	100.0	58	16	US-09-270-767-57408

ALIGNMENTS

RESULT 1
US-09-057-228-60
; Sequence 60, Application US/09057228
; GENERAL INFORMATION:
; APPLICANT: Yuriyeva, Olga
; APPLICANT: Kuriyan, John
; APPLICANT: O'Donnell, Michael E.
; APPLICANT: Jeruzalmi, David
; TITLE OF INVENTION: ENZYME DERIVED FROM THERMOPHILIC ORGANISMS THAT FUNCTION
; NUMBER OF SEQUENCES: 116
; CORRESPONDENCE ADDRESS:
; STREET: 411 Hackensack Ave, Continental Plaza, 4th
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/057,228
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/143,202

; FILING DATE: 08-APR-1997
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/823,407
 ; FILING DATE: 08-APR-1997
 ; CLASSIFICATION:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Jackson Esq., David A.
 ; REGISTRATION NUMBER: 26,742
 ; REFERENCE/DOCKET NUMBER: 600-1-179 N
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 201-487-5800
 ; TELEFAX: 201-343-1684
 ; INFORMATION FOR SEQ ID NO: 60:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 13 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; US-09-057-228-60

Query Match 100.0%; Score 23; DB 14; Length 13;
 Best Local Similarity 100.0%; Pred. No. 2.8e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DYDA 4
 Db 5 DYDA 8

RESULT 2
 US-09-057-416-60
 ; Sequence 60, Application US/09057416
 ; GENERAL INFORMATION:
 ; APPLICANT: Yuriyeva, Olga
 ; APPLICANT: Kuriyan, John
 ; APPLICANT: O'Donnell, Michael E.
 ; APPLICANT: Jeruzalmi, David
 ; TITLE OF INVENTION: ENZYME DERIVED FROM THERMOPHILIC ORGANISMS
 ; TITLE OF INVENTION: THAT FUNCTIONS AS A CHROMOSOMAL REPLICASE, PREPARATION AND USE
 ; NUMBER OF SEQUENCES: 116
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: David A. Jackson, Esq.
 ; STREET: 411 Hackensack Ave, Continental Plaza, 4th
 ; STREET: Floor
 ; CITY: Hackensack
 ; STATE: New Jersey
 ; COUNTRY: USA
 ; ZIP: 07601
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/057,416
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 60/143,202
 ; FILING DATE: 08-APR-1997
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/823,407
 ; FILING DATE: 08-APR-1997
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Jackson Esq., David A.
 ; REGISTRATION NUMBER: 26,742
 ; REFERENCE/DOCKET NUMBER: 600-1-179 N
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 201-487-5800

; TELEFAX: 201-343-1684
 ; INFORMATION FOR SEQ ID NO: 60:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 13 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; US-09-057-416-60

Query Match 100.0%; Score 23; DB 14; Length 13;
 Best Local Similarity 100.0%; Pred. No. 2.8e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DYDA 4
 Db 5 DYDA 8

RESULT 3
 US-09-716-964A-60
 ; Sequence 60, Application US/09716964A
 ; GENERAL INFORMATION:
 ; APPLICANT: O'Donnell, Michael E.
 ; APPLICANT: Yuzhakov, Alexander
 ; APPLICANT: Yuriyeva, Olga
 ; APPLICANT: Jeruzalmi, David
 ; APPLICANT: Bruck, Irina
 ; APPLICANT: Kuriyan, John
 ; TITLE OF INVENTION: ENZYMES DERIVED FROM THERMOPHILIC ORGANISMS THAT
 ; TITLE OF INVENTION: FUNCTION AS A CHROMOSOMAL REPLICASE, PREPARATION AND
 ; FILE REFERENCE: 22221/1030
 ; CURRENT APPLICATION NUMBER: US/09/716, 964A
 ; CURRENT FILING DATE: 2000-11-21
 ; PRIOR APPLICATION NUMBER: 60/143,202
 ; PRIOR FILING DATE: 1997-04-08
 ; PRIOR APPLICATION NUMBER: 08/823,407
 ; PRIOR FILING DATE: 1997-04-08
 ; PRIOR APPLICATION NUMBER: 09/057,416
 ; PRIOR FILING DATE: 1998-04-08
 ; NUMBER OF SEQ ID NOS: 212
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 60
 ; LENGTH: 13
 ; TYPE: PPT
 ; ORGANISM: Thermus thermophilus
 ; US-09-716-964A-60

Query Match 100.0%; Score 23; DB 21; Length 13;
 Best Local Similarity 100.0%; Pred. No. 2.8e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DYDA 4
 Db 5 DYDA 8

RESULT 4
 US-60-160-209-3652
 ; Sequence 3652, Application US/60160209
 ; GENERAL INFORMATION:
 ; APPLICANT: BONAZZI, VIVIAN
 ; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
 ; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS AND USES
 ; FILE REFERENCE: CLO00113
 ; CURRENT APPLICATION NUMBER: US/60/160,209
 ; CURRENT FILING DATE: 1999-10-19
 ; NUMBER OF SEQ ID NOS: 4646
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 3652
 ; LENGTH: 20

```
; TYPE: PRT
; ORGANISM: HUMAN
US-60-160-209-3652

Query Match      100.0%; Score 23; DB 27; Length 20;
Best Local Similarity 100.0%; Pred. No. 4.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DYDA 4
    ||||
Db 2 DYDA 5

RESULT 5
US-60-169-868-4801
; Sequence 4801, Application US/60169868
; GENERAL INFORMATION:
; APPLICANT: Bonazzi, Vivien
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL000161
; CURRENT APPLICATION NUMBER: US/60/169,868
; CURRENT FILING DATE: 1999-12-09
; NUMBER OF SEQ ID NOS: 6938
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4801
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Human
US-60-169-868-4801

Query Match      100.0%; Score 23; DB 27; Length 20;
Best Local Similarity 100.0%; Pred. No. 4.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DYDA 4
    ||||
Db 2 DYDA 5

RESULT 6
PCT-US00-09866A-10
; Sequence 10, Application PC/TUS0009866A
; GENERAL INFORMATION:
; APPLICANT: SMITHKLINE BEECHAM CORPORATION
; TITLE OF INVENTION: CONSERVED ADHESIN MOTIF AND METHODS OF
; TITLE OF INVENTION: USE THEREOF
; FILE REFERENCE: GM50047
; CURRENT APPLICATION NUMBER: PCT/US00/09866A
; CURRENT FILING DATE: 2000-04-13
; PRIOR APPLICATION NUMBER: US 60/129,073
; PRIOR FILING DATE: 1999-04-13
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Yersinia pestis
PCT-US00-09866A-10

Query Match      100.0%; Score 23; DB 1; Length 21;
Best Local Similarity 100.0%; Pred. No. 4.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DYDA 4
    ||||
Db 12 DYDA 15

RESULT 7
US-60-160-209-2607
; Sequence 2607, Application US/60160209
; GENERAL INFORMATION:
; APPLICANT: BONAZZI, VIVIAN
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS AND USES
; FILE REFERENCE: CL000113
; CURRENT APPLICATION NUMBER: US/60/160,209
; CURRENT FILING DATE: 1999-10-19
; NUMBER OF SEQ ID NOS: 4646
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2607
; LENGTH: 24
; TYPE: PRT
; ORGANISM: HUMAN
US-60-160-209-2607

Query Match      100.0%; Score 23; DB 27; Length 24;
Best Local Similarity 100.0%; Pred. No. 5.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DYDA 4
    ||||
Db 8 DYDA 11

RESULT 8
US-60-160-209-2726
; Sequence 2726, Application US/60160209
; GENERAL INFORMATION:
; APPLICANT: BONAZZI, VIVIAN
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS AND USES
; FILE REFERENCE: CL000113
; CURRENT APPLICATION NUMBER: US/60/160,209
; CURRENT FILING DATE: 1999-10-19
; NUMBER OF SEQ ID NOS: 4646
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2726
; LENGTH: 24
; TYPE: PRT
; ORGANISM: HUMAN
US-60-160-209-2726

Query Match      100.0%; Score 23; DB 27; Length 24;
Best Local Similarity 100.0%; Pred. No. 5.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DYDA 4
    ||||
Db 8 DYDA 11

RESULT 9
US-08-263-921B-36
; Sequence 36, Application US/08263921B
; GENERAL INFORMATION:
; APPLICANT: KEELING, PETER L
; APPLICANT: CHANG, MING-TANG
; TITLE OF INVENTION: NOVEL PLANTS AND PROCESSES FOR
; TITLE OF INVENTION: OBTAINING THEM
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: D S REMOLDT
; STREET: ICI SEEDS, 2369 330TH ST
; CITY: SLATER
; STATE: IA
; COUNTRY: UNITED STATES
; ZIP: 50244
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
```

;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/263,921B
;; FILING DATE: 21-JUN-1994
;; CLASSIFICATION: 800
;; ATTORNEY/AGENT INFORMATION:
;; NAME: REWOLDT, DANA S
;; REGISTRATION NUMBER: 33,762
;; REFERENCE/DOCKET NUMBER: 1.02.15
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (515)685-3574 EXT407
;; TELEFAX: (515) 685-2548
;; INFORMATION FOR SEQ ID NO: 36:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 28 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;; ORIGINAL SOURCE:
;; ORGANISM: PAGE 11 T4
;; US-08-263-921B-36

Query Match 100.0%; Score 23; DB 6; Length 28;
Best Local Similarity 100.0%; Pred. No. 6.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
|||
DB 17 DYDA 20

RESULT 10
US-08-346-602A-36
;; Sequence 36, Application US/08346602A
;; GENERAL INFORMATION:
;; APPLICANT: KEELING, PETER L
;; APPLICANT: CHANG, MING-TANG
;; TITLE OF INVENTION: NOVEL PLANTS AND PROCESSES FOR OBTAINING
;; TITLE OF INVENTION: THEM
;; NUMBER OF SEQUENCES: 43
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: DANA REWOLDT ICI SEEDS
;; STREET: 2369 330TH ST BOX 500
;; CITY: SLATER
;; STATE: IA
;; COUNTRY: US
;; ZIP: 50244
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/346,602A
;; FILING DATE: 29-NOV-1994
;; CLASSIFICATION: 800
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/263,921
;; FILING DATE: 21-JUN-1994
;; ATTORNEY/AGENT INFORMATION:
;; NAME: REWOLDT, DANA S
;; REGISTRATION NUMBER: 33,762
;; REFERENCE/DOCKET NUMBER: 1.02.15A
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (515) 685-3574
;; TELEFAX: (515) 685-2548
;; INFORMATION FOR SEQ ID NO: 36:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 28 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear

;; MOLECULE TYPE: peptide
;; ORIGINAL SOURCE:
;; ORGANISM: PAGE 11 T4
;; US-08-346-602A-36

Query Match 100.0%; Score 23; DB 7; Length 28;
Best Local Similarity 100.0%; Pred. No. 6.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
|||
DB 17 DYDA 20

RESULT 11
US-09-118-990A-36
;; Sequence 36, Application US/09118990A
;; GENERAL INFORMATION:
;; APPLICANT: KEELING, PETER L
;; APPLICANT: CHANG, MING-TANG
;; TITLE OF INVENTION: NOVEL PLANTS AND PROCESSES FOR OBTAINING
;; TITLE OF INVENTION: THEM
;; NUMBER OF SEQUENCES: 43
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: DANA REWOLDT ICI SEEDS
;; STREET: 2369 330TH ST BOX 500
;; CITY: SLATER
;; STATE: IA
;; COUNTRY: US
;; ZIP: 50244
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/118,990A
;; FILING DATE: 20-Jul-1998
;; CLASSIFICATION: <Unknown>
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/346,602
;; FILING DATE: 29-NOV-1994
;; APPLICATION NUMBER: US 08/263,921
;; FILING DATE: 21-JUN-1994
;; ATTORNEY/AGENT INFORMATION:
;; NAME: REWOLDT, DANA S
;; REGISTRATION NUMBER: 33,762
;; REFERENCE/DOCKET NUMBER: 1.02.15A
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (515) 685-3574
;; TELEFAX: (515) 685-2548
;; INFORMATION FOR SEQ ID NO: 36:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 28 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;; ORIGINAL SOURCE:
;; ORGANISM: PAGE 11 T4
;; SEQUENCE DESCRIPTION: SEQ ID NO: 36:
;; US-09-118-990A-36

Query Match 100.0%; Score 23; DB 15; Length 28;
Best Local Similarity 100.0%; Pred. No. 6.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
|||
DB 17 DYDA 20

RESULT 12

US-60-160-209-2539
; Sequence 2539, Application US/60160209
; GENERAL INFORMATION:
; APPLICANT: BONAZZI, VIVIEN
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS AND USES
; FILE REFERENCE: CLO00113
; CURRENT FILING DATE: 1999-10-19
; CURRENT APPLICATION NUMBER: US/60/160,209
; NUMBER OF SEQ ID NOS: 4646
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2539
; LENGTH: 28
; TYPE: PRT
; ORGANISM: HUMAN
US-60-160-209-2539

Query Match 100.0%; Score 23; DB 27; Length 28;
Best Local Similarity 100.0%; Pred. No. 6.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DYDA 4
|||||
Db 2 DYDA 5

RESULT 13
US-60-169-868-3774
; Sequence 3774, Application US/60169868
; GENERAL INFORMATION:
; APPLICANT: Bonazzi, Vivien
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CLO00161
; CURRENT APPLICATION NUMBER: US/60/169,868
; CURRENT FILING DATE: 1999-12-09
; NUMBER OF SEQ ID NOS: 6938
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3774
; LENGTH: 28
; TYPE: PRT
; ORGANISM: Human
US-60-169-868-3774

Query Match 100.0%; Score 23; DB 27; Length 28;
Best Local Similarity 100.0%; Pred. No. 6.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DYDA 4
|||||
Db 2 DYDA 5

RESULT 14
US-60-162-247-4129
; Sequence 4129, Application US/60162247
; GENERAL INFORMATION:
; APPLICANT: Bonazzi, Vivien
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS AND
; FILE REFERENCE: CLO00127
; CURRENT APPLICATION NUMBER: US/60/162,247
; CURRENT FILING DATE: 1999-10-29
; NUMBER OF SEQ ID NOS: 5442
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4129
; LENGTH: 31
; TYPE: PRT
; ORGANISM: Human
US-60-162-247-4129

Query Match 100.0%; Score 23; DB 27; Length 31;
Best Local Similarity 100.0%; Pred. No. 7.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DYDA 4
|||||
Db 17 DYDA 20

RESULT 15
PCT-US01-00663-36701
; Sequence 36701, Application PC/TUS0100663
; GENERAL INFORMATION:
; APPLICANT: Molecular Dynamics, Inc.
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL
; TITLE OF INVENTION: ANALYSIS OF GENE EXPRESSION IN HUMAN PLACENTA
; FILE REFERENCE: PB 0004 WO 7
; CURRENT APPLICATION NUMBER: PCT/US01/00663
; CURRENT FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 04 February 2000 (04.02.00)
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 26 May 2000 (26.05.00)
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 03 August 2000 (03.08.00)
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 03 October 2000 (03.10.00)
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 27 September 2000 (27.09.00)
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 21 September 2000 (21.09.00)
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 30 June 2000 (30.06.00)
; NUMBER OF SEQ ID NOS: 38837
; SOFTWARE: Molecular Dynamics Sequence Listing Engine
; SEQ ID NO 36701
; LENGTH: 38
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL157392.2
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.96
; OTHER INFORMATION: EST_HUMAN HIT: BF574717.1, EVALUATE 1.00e-14
; OTHER INFORMATION: SWISSPROT HIT: P08442, EVALUATE 7.40e-01
PCT-US01-00663-36701

Query Match 100.0%; Score 23; DB 1; Length 38;
Best Local Similarity 100.0%; Pred. No. 9.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DYDA 4
|||||
Db 19 DYDA 22

RESULT 16
US-08-417-872A-55
; Sequence 55, Application US/08417872A
; GENERAL INFORMATION:
; APPLICANT: SPARKS, ANDREW B
; APPLICANT: HOFFMAN, NOAH
; APPLICANT: KAY, BRIAN K.
; TITLE OF INVENTION: POLYPEPTIDES HAVING A FUNCTIONAL DOMAIN
; TITLE OF INVENTION: OF INTEREST AND METHODS OF IDENTIFYING AND USING SAME
; NUMBER OF SEQUENCES: 89
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas

CITY: New York
STATE: New York
COUNTRY: U.S.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/417,872A
FILING DATE: 07-APR-1995
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Mirock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 1101-193
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-790-9090
TELEFAX: 212-869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 55:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-417-872A-55

Query Match 100.0%; Score 23; DB 8; Length 38;
Best Local Similarity 100.0%; Pred. No. 9.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
|||
Db 5 DYDA 8

RESULT 17
US-09-864-761-45554
Sequence 45554, Application US/09864761
GENERAL INFORMATION:
APPLICANT: Penn, Sharron G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aeomica-x-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 2426316
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665

PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 45554
LENGTH: 38
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AL157392.2
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.96
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.83
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.74
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.88
OTHER INFORMATION: EST_HUMAN HIT: BF574717.1, EVALUE 1.00e-14
OTHER INFORMATION: SWISSPROT HIT: P08442, EVALUE 7.40e-01
US-09-864-761-45554

Query Match 100.0%; Score 23; DB 22; Length 38;
Best Local Similarity 100.0%; Pred. No. 9.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
|||
Db 19 DYDA 22

RESULT 18
US-09-879-957-102
Sequence 102, Application US/09879957
GENERAL INFORMATION:
APPLICANT: SPARKS, Andrew B.
HOFFMAN, Noah
KAY, Brian K.
FOWLKES, Dana M.
McCONNELL, Stephen J.
TITLE OF INVENTION: POLYPEPTIDES HAVING A FUNCTIONAL
DOMAIN OF INTEREST AND METHODS OF IDENTIFYING AND
USING SAME
NUMBER OF SEQUENCES: 227
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
SHEET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/879,957
FILING DATE: 13-Jun-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/630,915
FILING DATE: 03-APR-1996
ATTORNEY/AGENT INFORMATION:
NAME: Misrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 1101-174
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 102:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: unknown
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 102:
US-09-879-957-102

Query Match 100.0%; Score 23; DB 22; Length 38;
Best Local Similarity 100.0%; Pred. No. 9.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
|
|
|
|
Db 5 DYDA 8

RESULT 19
US-10-182-993-35614
; Sequence 35614, Application US/10182993
; GENERAL INFORMATION:
; APPLICANT: Molecular Dynamics, Inc.
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: ANALYSIS OF GENE EXPRESSION IN HUMAN BRAIN
; FILE REFERENCE: PB 0004 WO 2
; CURRENT APPLICATION NUMBER: US/10/182,993
; CURRENT FILING DATE: 2002-08-02
; PRIOR FILING DATE: US 60/180,312
; PRIOR FILING DATE: 04 February 2000 (04.02.00)
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 26 May 2000 (26.05.00)
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 03 August 2000 (03.08.00)
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 03 October 2000 (03.10.00)
; PRIOR FILING DATE: 27 September 2000 (27.09.00)
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 21 September 2000 (21.09.00)
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 30 June 2000 (30.06.00)
; NUMBER OF SEQ ID NOS: 37811
; SOFTWARE: Molecular Dynamics Sequence Listing Engine
; SEQ ID NO 35614
; LENGTH: 38
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO ALL157392.2
; FEATURE:
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.88
; FEATURE:
; OTHER INFORMATION: EST_HUMAN HIT: BF574717.1, EVALUE 1.00e-14
; FEATURE:
; OTHER INFORMATION: SWISSPROT HIT: P08442, EVALUE 7.40e-01
US-10-182-993-35614

Query Match 100.0%; Score 23; DB 25; Length 38;
Best Local Similarity 100.0%; Pred. No. 9.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
|
|
|
|
Db 19 DYDA 22

RESULT 20
US-10-203-134-36629
; Sequence 36629, Application US/10203134
; GENERAL INFORMATION:
; APPLICANT: Molecular Dynamics, Inc.
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL
; TITLE OF INVENTION: ANALYSIS OF GENE EXPRESSION IN HUMAN BONE MARROW
; FILE REFERENCE: PB 0004 WO 6
; CURRENT APPLICATION NUMBER: US/10/203,134
; CURRENT FILING DATE: 2002-08-02
; PRIOR FILING DATE: US 60/180,312
; PRIOR FILING DATE: 04 February 2000 (04.02.00)
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 26 May 2000 (26.05.00)
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 03 August 2000 (03.08.00)
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 03 October 2000 (03.10.00)
; PRIOR FILING DATE: 27 September 2000 (27.09.00)
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 21 September 2000 (21.09.00)
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 30 June 2000 (30.06.00)
; NUMBER OF SEQ ID NOS: 38628
; SOFTWARE: Molecular Dynamics Sequence Listing Engine
; SEQ ID NO 36629
; LENGTH: 38
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO ALL157392.2
; FEATURE:
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1
; FEATURE:
; OTHER INFORMATION: EST_HUMAN HIT: BF574717.1, EVALUE 1.00e-14
; FEATURE:
; OTHER INFORMATION: SWISSPROT HIT: P08442, EVALUE 7.40e-01
US-10-203-134-36629

Query Match 100.0%; Score 23; DB 26; Length 38;
Best Local Similarity 100.0%; Pred. No. 9.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
|
|
|
|
Db 19 DYDA 22

RESULT 21
US-10-203-136-36681
; Sequence 36681, Application US/10203136
; GENERAL INFORMATION:
; APPLICANT: Molecular Dynamics, Inc.
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL
; TITLE OF INVENTION: ANALYSIS OF GENE EXPRESSION IN HUMAN ADULT LIVER
US-10-203-136-36681

```
; FILE REFERENCE: PB 0004 WO 3
; CURRENT APPLICATION NUMBER: US/10/203,136
; CURRENT FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 04 February 2000 (04.02.00)
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 26 May 2000 (26.05.00)
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 03 August 2000 (03.08.00)
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 03 October 2000 (03.10.00)
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 27 September 2000 (27.09.00)
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 21 September 2000 (21.09.00)
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 30 June 2000 (30.06.00)
; NUMBER OF SEQ ID NOS: 38578
; SOFTWARE: Molecular Dynamics Sequence Listing Engine
; SEQ ID NO 36681
; LENGTH: 38
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL157392.2
; FEATURE:
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.83
; FEATURE:
; OTHER INFORMATION: EST_HUMAN HIT: BF574717.1, EVALUE 1.00e-14
; FEATURE:
; OTHER INFORMATION: SWISSPROT HIT: P08442, EVALUE 7.40e-01
; US-10-203-136-36681
```

```
Query Match 100.0%; Score 23; DB 26; Length 38;
Best Local Similarity 100.0%; Pred. No. 9.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 DYDA 4
    ||||
DB 19 DYDA 22
```

```
RESULT 22
US-10-203-137-36701
; Sequence 36701, Application US/10203137
; GENERAL INFORMATION:
; APPLICANT: Molecular Dynamics, Inc.
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: PB 0004 WO 7
; CURRENT APPLICATION NUMBER: US/10/203,137
; CURRENT FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 04 February 2000 (04.02.00)
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 26 May 2000 (26.05.00)
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 03 August 2000 (03.08.00)
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 03 October 2000 (03.10.00)
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 27 September 2000 (27.09.00)
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 21 September 2000 (21.09.00)
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 30 June 2000 (30.06.00)
; NUMBER OF SEQ ID NOS: 38837
; SOFTWARE: Molecular Dynamics Sequence Listing Engine
; SEQ ID NO 36701
```

```
; LENGTH: 38
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL157392.2
; FEATURE:
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.96
; FEATURE:
; OTHER INFORMATION: EST_HUMAN HIT: BF574717.1, EVALUE 1.00e-14
; FEATURE:
; OTHER INFORMATION: SWISSPROT HIT: P08442, EVALUE 7.40e-01
; US-10-203-137-36701
```

```
Query Match 100.0%; Score 23; DB 26; Length 38;
Best Local Similarity 100.0%; Pred. No. 9.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 DYDA 4
    ||||
DB 19 DYDA 22
```

```
RESULT 23
US-10-203-139-35253
; Sequence 35253, Application US/10203139
; GENERAL INFORMATION:
; APPLICANT: Molecular Dynamics, Inc.
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL
; FILE REFERENCE: PB 0004 WO 4
; CURRENT APPLICATION NUMBER: US/10/203,139
; CURRENT FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 04 February 2000 (04.02.00)
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 26 May 2000 (26.05.00)
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 03 August 2000 (03.08.00)
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 03 October 2000 (03.10.00)
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 27 September 2000 (27.09.00)
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 21 September 2000 (21.09.00)
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 30 June 2000 (30.06.00)
; NUMBER OF SEQ ID NOS: 37156
; SOFTWARE: Molecular Dynamics Sequence Listing Engine
; SEQ ID NO 35253
; LENGTH: 38
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL157392.2
; FEATURE:
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.74
; FEATURE:
; OTHER INFORMATION: EST_HUMAN HIT: BF574717.1, EVALUE 1.00e-14
; FEATURE:
; OTHER INFORMATION: SWISSPROT HIT: P08442, EVALUE 7.40e-01
; US-10-203-139-35253
```

```
Query Match 100.0%; Score 23; DB 26; Length 38;
Best Local Similarity 100.0%; Pred. No. 9.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 DYDA 4
    ||||
DB 19 DYDA 22
```

RESULT 24
US-60-161-932-1951
; Sequence 1951, Application US/60161932
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; TITLE OF INVENTION: The Drosophila Genome, Primary Nucleic
; TITLE OF INVENTION: Acid Sequences, Systems Containin the Nucleic Acid Sequences
; FILE REFERENCE: CL000122
; CURRENT APPLICATION NUMBER: US/60/161.932
; CURRENT FILING DATE: 1999-10-28
; NUMBER OF SEQ ID NOS: 2626
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1951
; LENGTH: 38
; TYPE: PRT
; ORGANISM: Drosophila
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(38)
; OTHER INFORMATION: Xaa - Any Amino Acid
US-60-161-932-1951

Query Match 100.0%; Score 23; DB 27; Length 38;
Best Local Similarity 100.0%; Pred. No. 9.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DYDA 4
Db 29 DYDA 32

RESULT 25
US-08-417-872A-57
; Sequence 57, Application US/08417872A
; GENERAL INFORMATION:
; APPLICANT: SPARKS, ANDREW B
; APPLICANT: HOFFMAN, NOAH
; APPLICANT: KAY, BRIAN K.
; TITLE OF INVENTION: POLYPEPTIDES HAVING A FUNCTIONAL DOMAIN
; TITLE OF INVENTION: OF INTEREST AND METHODS OF IDENTIFYING AND USING SAME
; NUMBER OF SEQUENCES: 89
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/417.872A
; FILING DATE: 07-APR-1995
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Misrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 1101-193
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 57:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 39 amino acids
; TYPE: amino acid
; STRANDEDNESS: single

; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-417-872A-57

Query Match 100.0%; Score 23; DB 8; Length 39;
Best Local Similarity 100.0%; Pred. No. 9.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DYDA 4
Db 5 DYDA 8

RESULT 26
US-09-879-957-104
; Sequence 104, Application US/09879957
; GENERAL INFORMATION:
; APPLICANT: SPARKS, Andrew B.
; APPLICANT: HOFFMAN, Noah
; KAY, Brian K.
; FOWLKES, Dana M.
; MCCONNELL, Stephen J.
; TITLE OF INVENTION: POLYPEPTIDES HAVING A FUNCTIONAL
; DOMAIN OF INTEREST AND METHODS OF IDENTIFYING AND
; USING SAME
; NUMBER OF SEQUENCES: 227
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/879.957
; FILING DATE: 13-Jun-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/630,915
; FILING DATE: 03-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Misrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 1101-174
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 104:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 39 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 104:
US-09-879-957-104

Query Match 100.0%; Score 23; DB 22; Length 39;
Best Local Similarity 100.0%; Pred. No. 9.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DYDA 4
Db 5 DYDA 8

RESULT 27
US-08-417-872A-49
; Sequence 49, Application US/08417872A
; GENERAL INFORMATION:
; APPLICANT: SPARKS, ANDREW B
; APPLICANT: HOFFMAN, NOAH
; APPLICANT: KAY, BRIAN K.
; TITLE OF INVENTION: POLYPEPTIDES HAVING A FUNCTIONAL DOMAIN
; OF INTEREST AND METHODS OF IDENTIFYING AND USING SAME
; NUMBER OF SEQUENCES: 89
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/08/417,872A
; FILING DATE: 07-APR-1995
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Mistrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 1101-193
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 49:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 40 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-417-872A-49

Query Match 100.0%; Score 23; DB 8; Length 40;
Best Local Similarity 100.0%; Pred. No. 1e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
Db 5 DYDA 8

RESULT 28
US-09-879-957-96
; Sequence 96, Application US/09879957
; GENERAL INFORMATION:
; APPLICANT: SPARKS, Andrew B.
; APPLICANT: HOFFMAN, Noah
; APPLICANT: KAY, Brian K.
; APPLICANT: FOWLKES, Dana M.
; APPLICANT: MCCONNELL, Stephen J.
; TITLE OF INVENTION: POLYPEPTIDES HAVING A FUNCTIONAL
; DOMAIN OF INTEREST AND METHODS OF IDENTIFYING AND
; USING SAME
; NUMBER OF SEQUENCES: 227
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York

; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/879,957
; FILING DATE: 13-Jun-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/630,915
; FILING DATE: 03-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Mistrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 1101-174
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 96:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 40 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 96:
US-09-879-957-96

Query Match 100.0%; Score 23; DB 22; Length 40;
Best Local Similarity 100.0%; Pred. No. 1e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
Db 5 DYDA 8

RESULT 29
US-09-513-996A-9365
; Sequence 9365, Application US/09513996A
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEP
; TITLE OF INVENTION: ENCODED THEREBY
; FILE REFERENCE: 2750-709p
; CURRENT APPLICATION NUMBER: US/09/513,996A
; CURRENT FILING DATE: 2000-02-25
; NUMBER OF SEQ ID NOS: 81028
; SEQ ID NO 9365
; LENGTH: 41
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 1..41 /
; OTHER INFORMATION: any n or Xaa = unknown
; FEATURE:
; OTHER INFORMATION: Location 1..41 / Ceres Seq. ID 1030234
US-09-513-996A-9365

Query Match 100.0%; Score 23; DB 19; Length 41;
Best Local Similarity 100.0%; Pred. No. 1e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
Db 28 DYDA 31

```

RESULT 30
US-09-513-996A-68915
; Sequence 68915, Application US/09513996A
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-709P
; CURRENT APPLICATION NUMBER: US/09/513,996A
; CURRENT FILING DATE: 2000-02-25
; NUMBER OF SEQ ID NOS: 81028
; SEQ ID NO 68915
; LENGTH: 41
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; NAME/KEY: UNSURE
; LOCATION: 1..41 /
; OTHER INFORMATION: any n or Xaa = unknown
; FEATURE:
; OTHER INFORMATION: Location 1..41 / Ceres Seq. ID 2185473
US-09-513-996A-68915

Query Match      100.0%; Score 23; DB 19; Length 41;
Best Local Similarity 100.0%; Pred. No. 1e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
Db 28 DYDA 31

RESULT 31
US-09-620-394B-5748
; Sequence 5748, Application US/09620394B
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai
; APPLICANT: BROVER, Vyacheslav
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1067P
; CURRENT APPLICATION NUMBER: US/09/620,394B
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 9131
; SEQ ID NO 5748
; LENGTH: 41
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; NAME/KEY: misc_feature
; LOCATION: 1..41
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: 1..41
; OTHER INFORMATION: Ceres Seq. ID 1394958
US-09-620-394B-5748

Query Match      100.0%; Score 23; DB 20; Length 41;
Best Local Similarity 100.0%; Pred. No. 1e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
Db 28 DYDA 31

RESULT 32
US-08-247-491-3
; Sequence 3, Application US/08247491
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: Yother, Janet L
; APPLICANT: McDaniel, Larry S

```

```

; TITLE OF INVENTION: STRUCTURAL GENE OF PNEUMOCOCCAL PROTEIN
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Shoemaker and Mattare, Ltd
; STREET: Suite 1203, 2001 Jefferson Davis Highway
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202-0286
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/247,491
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/835,698
; FILING DATE: 12-FEB-1992
; APPLICATION NUMBER: US 07/656,773
; FILING DATE: 15-FEB-1991
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 415-0810
; TELEFAX: (703) 521-0378
; TELEX: LUKPAT WASHINGTON
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 45 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-247-491-3

Query Match      100.0%; Score 23; DB 6; Length 45;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
Db 14 DYDA 17

RESULT 33
US-08-467-852-3
; Sequence 3, Application US/08467852
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: Yother, Janet L
; APPLICANT: McDaniel, Larry S
; TITLE OF INVENTION: STRUCTURAL GENE OF PNEUMOCOCCAL PROTEIN
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Shoemaker and Mattare, Ltd
; STREET: Suite 1203, 2001 Jefferson Davis Highway
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202-0286
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/467,852
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/247,491
; FILING DATE: 23-MAY-1994
; APPLICATION NUMBER: US 07/835,698
; FILING DATE: 12-FEB-1992

```

```

; APPLICATION NUMBER: US 07/656,773
; FILING DATE: 15-FEB-1991
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 415-0810
; TELEFAX: (703) 521-0378
; TELEX: LUKPAT WASHINGTON
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 45 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-467-852-3

Query Match 100.0%; Score 23; DB 8; Length 45;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
Db 14 DYDA 17

RESULT 34
US-08-470-626-3
; Sequence 3, Application US/08470626,
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: Yother, Janet L
; APPLICANT: McDaniel, Larry S
; TITLE OF INVENTION: STRUCTURAL GENE OF PNEUMOCOCCAL PROTEIN
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Shoemaker and Mattare, Ltd
; STREET: Suite 1203, 2001 Jefferson Davis Highway
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202-0286
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/470,626
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/835,698
; FILING DATE: 12-FEB-1992
; APPLICATION NUMBER: US 07/656,773
; FILING DATE: 15-FEB-1991
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 415-0810
; TELEFAX: (703) 521-0378
; TELEX: LUKPAT WASHINGTON
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 45 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-470-626-3

Query Match 100.0%; Score 23; DB 8; Length 45;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
Db 14 DYDA 17

RESULT 35
US-08-470-626-3
; Sequence 3, Application US/08470626,
; GENERAL INFORMATION:
; APPLICANT: Briles, David E
; APPLICANT: Yother, Janet L
; APPLICANT: McDaniel, Larry S
; TITLE OF INVENTION: STRUCTURAL GENE OF PNEUMOCOCCAL PROTEIN
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Shoemaker and Mattare, Ltd
; STREET: Suite 1203, 2001 Jefferson Davis Highway
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202-0286
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/470,626
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/835,698
; FILING DATE: 12-FEB-1992
; APPLICATION NUMBER: US 07/656,773
; FILING DATE: 15-FEB-1991
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 415-0810
; TELEFAX: (703) 521-0378
; TELEX: LUKPAT WASHINGTON
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 45 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-470-626-3

Query Match 100.0%; Score 23; DB 8; Length 45;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
Db 14 DYDA 17

RESULT 36
US-09-450-969-3984
; Sequence 3984, Application US/09450969
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: PATH99-09A
; CURRENT APPLICATION NUMBER: US/09/450,969
; CURRENT FILING DATE: 1999-11-29
; NUMBER OF SEQ ID NOS: 7544
; SEQ ID NO 3984
; LENGTH: 46
; TYPE: PRT
; ORGANISM: S.epidermidis
US-09-450-969-3984

Query Match 100.0%; Score 23; DB 18; Length 46;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
Db 13 DYDA 16

RESULT 37
US-10-008-524A-154
; Sequence 154, Application US/10008524A
; GENERAL INFORMATION:
; APPLICANT: Doorbar, John
; TITLE OF INVENTION: IMPROVEMENTS IN OR RELATING TO SCREENING FOR PAPILLOMAVIRUS
; FILE REFERENCE: 18396/1074
; CURRENT APPLICATION NUMBER: US/10/008,524A
; CURRENT FILING DATE: 2002-08-13
; PRIOR APPLICATION NUMBER: 09/314,268
; PRIOR FILING DATE: 1999-05-18
; NUMBER OF SEQ ID NOS: 179
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 154
; LENGTH: 47
; TYPE: PRT
; ORGANISM: Human papillomavirus type 60
US-10-008-524A-154
```

```

US-60-163-062-951
; Sequence 951, Application US/60163062
; GENERAL INFORMATION:
; APPLICANT: Bonazzi, Vivien
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN KINASE PROTEINS AND USES THEREOF
; FILE REFERENCE: CL000134
; CURRENT APPLICATION NUMBER: US/60/163,062
; CURRENT FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 1302
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 951
; LENGTH: 45
; TYPE: PRT
; ORGANISM: Human
US-60-163-062-951

Query Match 100.0%; Score 23; DB 27; Length 45;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
Db 6 DYDA 9

RESULT 38
US-09-450-969-3984
; Sequence 3984, Application US/09450969
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: PATH99-09A
; CURRENT APPLICATION NUMBER: US/09/450,969
; CURRENT FILING DATE: 1999-11-29
; NUMBER OF SEQ ID NOS: 7544
; SEQ ID NO 3984
; LENGTH: 46
; TYPE: PRT
; ORGANISM: S.epidermidis
US-09-450-969-3984

Query Match 100.0%; Score 23; DB 18; Length 46;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
Db 13 DYDA 16

RESULT 39
US-10-008-524A-154
; Sequence 154, Application US/10008524A
; GENERAL INFORMATION:
; APPLICANT: Doorbar, John
; TITLE OF INVENTION: IMPROVEMENTS IN OR RELATING TO SCREENING FOR PAPILLOMAVIRUS
; FILE REFERENCE: 18396/1074
; CURRENT APPLICATION NUMBER: US/10/008,524A
; CURRENT FILING DATE: 2002-08-13
; PRIOR APPLICATION NUMBER: 09/314,268
; PRIOR FILING DATE: 1999-05-18
; NUMBER OF SEQ ID NOS: 179
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 154
; LENGTH: 47
; TYPE: PRT
; ORGANISM: Human papillomavirus type 60
US-10-008-524A-154
```


Query Match 100.0%; Score 23; DB 24; Length 47;
 Best Local Similarity 100.0%; Pred. No. 1.2e+03;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
 DB 23 DYDA 26

RESULT 38
 US-08-651-136A-58
 ; Sequence 58, Application US/08651136A
 ; GENERAL INFORMATION:
 ; APPLICANT: Schulein, Martin
 ; APPLICANT: Andersen, Lene N.
 ; APPLICANT: Lassen, Soren F.
 ; APPLICANT: Kauppinen, Markus S.
 ; APPLICANT: Lange, Lene
 ; APPLICANT: Nielsen, Ruby I.
 ; APPLICANT: Ihara, Michiko
 ; APPLICANT: Takagi, Shinobu
 ; TITLE OF INVENTION: Novel Endoglucanases
 ; NUMBER OF SEQUENCES: 101
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Novo Nordisk of North America, Inc.
 ; STREET: 405 Lexington Avenue, 64th Floor
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: United States of America
 ; ZIP: 10174-6401
 ; COMPUTER READABLE FORM:
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION NUMBER: US/08/651,136A
 ; FILING DATE: 21-MAY-1996
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Lambiris, Elias J.
 ; REGISTRATION NUMBER: 33,728
 ; REFERENCE/DOCKET NUMBER: 4366.200-US
 ; TELEPHONE: 212-867-0123
 ; TELEFAX: 212-878-9655
 ; INFORMATION FOR SEQ ID NO: 58:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 49 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-651-136A-58

Query Match 100.0%; Score 23; DB 10; Length 49;
 Best Local Similarity 100.0%; Pred. No. 1.3e+03;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
 DB 11 DYDA 14

RESULT 39
 US-10-007-521-58
 ; Sequence 58, Application US/10007521
 ; GENERAL INFORMATION:
 ; APPLICANT: Schulein, Martin
 ; APPLICANT: Andersen, Lene N.
 ; APPLICANT: Lassen, Soren F.
 ; APPLICANT: Kauppinen, Markus S.
 ; APPLICANT: Lange, Lene
 ; APPLICANT: Nielsen, Ruby I.

Ihara, Michiko
 Takagi, Shinobu
 TITLE OF INVENTION: Novel Endoglucanases
 NUMBER OF SEQUENCES: 109
 CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Novo Nordisk of North America, Inc.
 ; STREET: 405 Lexington Avenue, 64th Floor
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: United States of America
 ; ZIP: 10174-6401
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/10/007,521
 ; FILING DATE: 10-Dec-2001
 ; CLASSIFICATION: <Unknown>
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/651,136
 ; FILING DATE: 21-MAY-1996
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Lambiris, Elias J.
 ; REGISTRATION NUMBER: 33,728
 ; REFERENCE/DOCKET NUMBER: 4366.200-US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 212-867-0123
 ; TELEFAX: 212-878-9655
 ; INFORMATION FOR SEQ ID NO: 58:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 49 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 58:
 US-10-007-521-58

Query Match 100.0%; Score 23; DB 24; Length 49;
 Best Local Similarity 100.0%; Pred. No. 1.3e+03;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
 DB 11 DYDA 14

RESULT 40
 US-60-177-571-4760
 ; Sequence 4760, Application US/60177571
 ; GENERAL INFORMATION:
 ; APPLICANT: Bonazzi, Vivien
 ; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
 ; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS,
 ; TITLE OF INVENTION: AND USES THEREOF
 ; FILE REFERENCE: CLO00201
 ; CURRENT APPLICATION NUMBER: US/60/177,571
 ; CURRENT FILING DATE: 2000-01-27
 ; NUMBER OF SEQ ID NOS: 5082
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 4760
 ; LENGTH: 51
 ; TYPE: PRT
 ; ORGANISM: HUMAN
 ; FEATURE:
 ; NAME/KEY: VARIANT
 ; LOCATION: (1)...(51)
 ; OTHER INFORMATION: xaa = Any Amino Acid
 US-60-177-571-4760
 Query Match 100.0%; Score 23; DB 27; Length 51;
 Best Local Similarity 100.0%; Pred. No. 1.3e+03;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
|||||

Db 22 DYDA 25

RESULT 41
US-10-015-127-10406
; Sequence 10406, Application US/10015127
; GENERAL INFORMATION:
; APPLICANT: Bower, Stanley G.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; TITLE OF INVENTION: Spingomonas elodea genome sequences and uses thereof
; FILE REFERENCE: 38-10(15806)B
; CURRENT APPLICATION NUMBER: US/10/015,127
; CURRENT FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: US 60/252,455
; PRIOR FILING DATE: 2000-11-22
; NUMBER OF SEQ ID NOS: 14357
; SEQ ID NO 10406
; LENGTH: 52
; TYPE: PRT
; ORGANISM: Spingomonas elodea
US-10-015-127-10406

Query Match 100.0%; Score 23; DB 24; Length 52;
Best Local Similarity 100.0%; Pred. No. 1.4e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
|||||

Db 45 DYDA 48

RESULT 42
US-60-360-039-12882
; Sequence 12882, Application US/60360039
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Chen, Xianfeng
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)A
; CURRENT APPLICATION NUMBER: US/60/360,039
; CURRENT FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 12882
; LENGTH: 53
; TYPE: PRT
; ORGANISM: Aspergillus nidulans
US-60-360-039-12882

Query Match 100.0%; Score 23; DB 27; Length 53;
Best Local Similarity 100.0%; Pred. No. 1.4e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
|||||

Db 5 DYDA 8

RESULT 43
US-10-029-386-28441
; Sequence 28441, Application US/10029386
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.

; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 28441
; LENGTH: 54
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR3.1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.2
; OTHER INFORMATION: SWISSPROT HIT: P51178, EVALU 9.00e-25
US-10-029-386-28441

Query Match 100.0%; Score 23; DB 24; Length 54;
Best Local Similarity 100.0%; Pred. No. 1.4e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
|||||

Db 29 DYDA 32

RESULT 44
PCT-US01-03800A-1476
; Sequence 1476, Application PC/TUS0103800A
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-030
; CURRENT APPLICATION NUMBER: PCT/US01/03800A
; CURRENT FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 2700
; SOFTWARE: Custom
; SEQ ID NO 1476
; LENGTH: 58
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(58)
; OTHER INFORMATION: Xaa = any amino acid or nothing
PCT-US01-03800A-1476

Query Match 100.0%; Score 23; DB 1; Length 58;
Best Local Similarity 100.0%; Pred. No. 1.6e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
|||||

Db 20 DYDA 23

RESULT 45
US-09-270-767-57408
; Sequence 57408, Application US/09270767
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.

; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 57408
; LENGTH: 58
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-57408

Query Match 100.0%; Score 23; DB 16; Length 58;
Best Local Similarity 100.0%; Pred. No. 1.6e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
|
|
|
|
Db 7 DYDA 10

Search completed: February 6, 2003, 11:37:30
Job time : 136.167 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 6, 2003, 11:16:04 ; Search time 10.5 Seconds
(without alignments)
36.623 Million cell updates/sec

Title: PAT943-1

Perfect score: 23

Sequence: 1 dyda 4

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 757

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 100%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR_73:*

2: pir1:*

3: pir2:*

4: pir3:*

5: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	23	100.0	44	2 S77832	probable transcrip
2	23	100.0	56	2 JW0095	p5 protein - human
3	23	100.0	67	2 D87301	hypothetical prote
4	23	100.0	67	2 D30540	conserved hypothet
5	23	100.0	68	2 AB1425	hypothetical prote
6	23	100.0	85	2 AD3215	hypothetical prote
7	23	100.0	86	2 H70721	hypothetical 10.2K
8	23	100.0	86	2 JH0445	hypothetical prote
9	23	100.0	90	2 A96968	hypothetical prote
10	23	100.0	91	2 S03306	ig heavy chain V r
11	23	100.0	91	2 E75531	ribosomal protein
12	23	100.0	93	2 D70970	hypothetical prote
13	23	100.0	93	2 F90686	hypothetical prote
14	23	100.0	93	2 B85537	hypothetical prote
15	23	100.0	95	2 A10048	hypothetical prote
16	23	100.0	100	2 F70530	hypothetical prote
17	23	100.0	100	2 D87037	conserved hypothet
18	23	100.0	102	2 T27821	hypothetical prote
19	23	100.0	103	1 WBP22	gene 17 protein -
20	23	100.0	103	2 S24187	aspartic proteinas
21	23	100.0	105	2 A1985	nitrogen fixation
22	23	100.0	105	2 G59903	hypothetical prote
23	23	100.0	107	2 AG3321	hypothetical prote
24	23	100.0	109	2 T01654	ribosomal protein
25	23	100.0	109	2 H87435	transcription regu
26	23	100.0	109	2 C84357	hypothetical prote
27	23	100.0	112	2 S19976	ig kappa chain V r
28	23	100.0	112	2 T12731	hypothetical prote
29	23	100.0	112	2 F87193	probable secreted
					hypothetical prote

ALIGNMENTS

RESULT 1

S77832

probable transcription elongation factor - Mycoplasma capricolum (fragment)
N:Alternate names: protein MC277
C:Species: Mycoplasma capricolum
C:Date: 09-Oct-1997 #sequence_revision 07-Nov-1997 #text_change 20-Jun-2000
C:Accession: S77832; S48616
R:Bork, P.; Ouzounis, C.; Casari, G.; Schneider, R.; Sander, C.; Dolan, M.; Gilbe
Mol. Microbiol. 16, 955-967, 1995
A:Title: Exploring the Mycoplasma capricolum genome: a minimal cell reveals its p
A:Reference number: S77739; MUID:96059641; PMID:7476192
A:Accession: S77832
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-44 <BOR>
A:Cross-references: EMBL:Z33203; NID:9541713; PIDN:CAA83788.1; PID:9950075
A:Experimental source: ATCC 27343
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, July 1994

C:Genetics:

A:Genetic code: SGC3

C:Superfamily: transcription elongation factor greb

C:Keywords: transcription factor

Query Match 100.0%; Score 23; DB 2; Length 44;

Best Local Similarity 100.0%; Pred. NO. 67;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4

|||||

Db 29 DYDA 32

RESULT 2

JW0095

p5 protein - human (fragments)

C:Species: Homo sapiens (man)

C:Date: 03-Jul-1998 #sequence_revision 10-Jul-1998 #text_change 10-Jul-1998

C:Accession: JW0095

R:Harrington, M.G.; Aebersold, R.; Martin, B.M.; Merrill, C.R.; Hood, L.

Appl. Theor. Electrophor. 3, 229-234, 1993

A:Title: Identification of a brain-specific human cerebrospinal fluid glycoprotein

A:Reference number: JW0095; MUID:94032630; PMID:7692978

A:Accession: JW0095

A:Status: preliminary

A:Molecule type: protein

A:Residues: 1-56 <HAR>

Query Match 100.0%; Score 23; DB 2; Length 56;

Best Local Similarity 100.0%; Pred. NO. 87;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4

Db 8 DYDA 11
||||
RESULT 3
D98266
hypothetical protein AGR_L2160 [imported] - Agrobacterium tumefaciens (strain C58, Cere
C:Species: Agrobacterium tumefaciens
C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 11-Jan-2002
C:Accession: D98266
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorollo, B.; Goldman,
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A:Reference number: A97359; PMID:11743194
A:Accession: D98266
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-67 <KUR>
A:Cross-references: GB:AE007870; PIDN:AAK89654.1; PID:g15159554; GSPDB:GN00170
C:Genetics:
A:Gene: AGR_L2160
A:Map position: linear chromosome
Query Match 100.0%; Score 23; DB 2; Length 67;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DYDA 4
||||
Db 22 DYDA 25
RESULT 4
AE3018
conserved hypothetical protein Atu3753 [imported] - Agrobacterium tumefaciens (strain C5
C:Species: Agrobacterium tumefaciens
C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 11-Jan-2002
C:Accession: AE3018
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClell
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; PMID:11743193
A:Accession: AE3018
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-67 <KUR>
A:Cross-references: GB:AE008689; PIDN:AAK44563.1; PID:g17742179; GSPDB:GN00187
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: Atu3753
A:Map position: linear chromosome
Query Match 100.0%; Score 23; DB 2; Length 67;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DYDA 4
||||
Db 22 DYDA 25
RESULT 5
AD3215
hypothetical protein Atu5451 [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C:Species: Agrobacterium tumefaciens
C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 11-Jan-2002
C:Accession: AD3215
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I

erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.;
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon
ster, E.W.
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58
A:Reference number: AB2577; PMID:11743193
A:Accession: AD3215
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-68 <KUR>
A:Cross-references: GB:AE008687; PIDN:AAK46138.1; PID:g17743905; GSPDB:GN00188
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: Atu5451
A:Genome: plasmid
Query Match 100.0%; Score 23; DB 2; Length 68;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DYDA 4
||||
Db 33 DYDA 36
RESULT 6
H70721
hypothetical protein Rv1831 - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
C:Accession: H70721
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Go
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holr
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: H70721
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-85 <COL>
A:Cross-references: GB:78020; GB:AL123456; NID:g3261625; PIDN:CAR01471.1; PID:625
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: Rv1831
Query Match 100.0%; Score 23; DB 2; Length 85;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DYDA 4
||||
Db 46 DYDA 49
RESULT 7
JH0445
hypothetical 10.2K protein (lyc 5' region) - Clostridium acetobutylicum
C:Species: Clostridium acetobutylicum
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 05-Jan-1996
C:Accession: JH0445
R:Croux, C.; Garcia, J.L.
Gene 104, 25-31, 1991
A:Title: Sequence of the lyc gene encoding the autolytic lysozyme of Clostridium a
A:Reference number: JH0441; MUID:92009192; PMID:1916274
A:Accession: JH0445
A:Molecule type: DNA
A:Residues: 1-86 <CRO>
A:Cross-references: GB:M68865
A:Experimental source: ATCC 824

Query Match 100.0%; Score 23; DB 2; Length 86;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
 ||||
 DB 4 DYDA 7

RESULT 8

A96968
 hypothetical protein CAC0553 [imported] - Clostridium acetobutylicum
 C:Species: Clostridium acetobutylicum
 C:Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 14-Sep-2001
 C:Accession: A96968
 R:Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee, J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
 J. Bacteriol. 183, 4823-4836, 2001
 A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Cld
 A:Reference number: A96900; MUID:21359325; PMID:21359325
 A:Accession: A96968
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-86 <KOR>
 A:Cross-references: GB:AE001437; PIDN:AAK78532.1; PID:g15023419; GSPDB:GN00168
 A:Experimental source: Clostridium acetobutylicum ATCC824
 C:Genetics:
 A:Gene: CAC0553

Query Match 100.0%; Score 23; DB 2; Length 86;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
 ||||
 DB 4 DYDA 7

RESULT 9

S03306
 Ig heavy chain V region (61B8) - mouse (fragment)
 C:Species: Mus musculus (house mouse)
 C:Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 23-Jul-1999
 C:Accession: J00046; S03306
 R:Van Cleave, V.H.; Naeye, C.W.; Metzger, D.W.
 J. Exp. Med. 167, 1841-1848, 1988
 A:Title: Do antibodies recognize amino acid side chains of protein antigens independently
 A:Reference number: J00043; MUID:88258372; PMID:2455014
 A:Accession: J00046
 A:Molecule type: mRNA
 A:Residues: 1-90 <VAN>
 A:Cross-references: EMBL:X12382; NID:g52095; PIDN:CAA30940.1; PID:g930171
 A:Superfamily: immunoglobulin V region, immunoglobulin homology
 C:Keywords: heterotetramer; immunoglobulin

Query Match 100.0%; Score 23; DB 2; Length 90;
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
 ||||
 DB 76 DYDA 79

RESULT 10

E75531
 ribosomal protein S15 - Deinococcus radiodurans (strain R1)
 C:Species: Deinococcus radiodurans
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 19-May-2000
 C:Accession: E75531
 R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma, S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.

Science 286, 1571-1577, 1999

A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1
 A:Reference number: A75250; MUID:20036896; PMID:10567266

A:Accession: E75531
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-91 <WHI>
 A:Cross-references: GB:AE001894; GB:AE000513; NID:g6458011; PIDN:AAF09921.1; PID:
 A:Experimental source: strain R1
 C:Genetics:

A:Gene: DR0341

A:Map position: 1

C:Superfamily: Escherichia coli ribosomal protein S15; eubacterial ribosomal prot

Query Match 100.0%; Score 23; DB 2; Length 91;
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
 ||||
 DB 70 DYDA 73

RESULT 11

D70970
 hypothetical protein RV3357 - Mycobacterium tuberculosis (strain H37RV)
 C:Species: Mycobacterium tuberculosis
 C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
 C:Accession: D70970
 R:Colle, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Go
 Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Ho
 Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
 Nature 393, 537-544, 1998
 A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
 A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete
 A:Reference number: A70500; MUID:98295987; PMID:9634230
 A:Accession: D70970
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-91 <COL>
 A:Cross-references: GB:AL009198; GB:AL123456; NID:g3242262; PIDN:CAAL5742.1; PID:
 A:Experimental source: strain H37RV
 C:Genetics:
 A:Gene: RV3357

Query Match 100.0%; Score 23; DB 2; Length 91;
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
 ||||
 DB 42 DYDA 45

RESULT 12

F90686
 hypothetical protein ECs0462 [imported] - Escherichia coli (strain O157:H7, subst
 C:Species: Escherichia coli
 C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 18-Jul-2001
 C:Accession: F90686
 R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Ha
 gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
 DNA Res. 8, 11-22, 2001
 A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and
 A:Reference number: A99629; MUID:21156231; PMID:11258796
 A:Accession: F90686
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-93 <HAY>

A:Cross-references: GB:BA000007; PIDN:BA833885.1; PID:g13359919; GSPDB:GN00154
 A:Experimental source: strain O157:H7, substrain RMD 0509952
 C:Genetics:
 A:Gene: ECs0462

Query Match 100.0%; Score 23; DB 2; Length 93;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
||||
Db 51 DYDA 54

RESULT 13

B85537
hypothetical protein ECs0462 [imported] - Escherichia coli (strain O157:H7, substrain EDL933)
C:Species: Escherichia coli
C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 27-Nov-2001
C:Accession: B85537

R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
Miller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001

A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.

A:Reference number: A85480; MUID:21074935; PMID:11206551

A:Accession: B85537

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-93 <STO>

A:Cross-references: GB:AF005174; NID:g12513261; PIDN:AAG54758.1; GSPDB:GN00145; UWGP:Z05

A:Experimental source: strain O157:H7, substrain EDL933

C:Genetics:

A:Gene: Z0510

Query Match 100.0%; Score 23; DB 2; Length 93;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
||||
Db 51 DYDA 54

RESULT 14

AI0048
hypothetical protein YPO0394 [imported] - Yersinia pestis (strain CO92)

C:Species: Yersinia pestis

C>Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 02-Nov-2001

C:Accession: AI0048

R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.
deno-Farraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,
Nature 413, 523-527, 2001

A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.

A:Reference number: AB0001; MUID:21470413; PMID:11586360

A:Accession: AI0048

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-95 <KUR>

A:Cross-references: GB:AL590842; PIDN:CAC89252.1; PID:g15978491; GSPDB:GN00175

C:Genetics:

A:Gene: YPO0394

Query Match 100.0%; Score 23; DB 2; Length 95;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
||||
Db 33 DYDA 36

RESULT 15

F70530
hypothetical protein Rv2699c - Mycobacterium tuberculosis (strain H37Rv)

C:Species: Mycobacterium tuberculosis

C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000

C:Accession: F70530

R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; G.
Connor, R.T.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holt
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998

A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete
A:Reference number: A70500; MUID:98295987; PMID:9634230

A:Accession: F70530

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-100 <COL>

A:Cross-references: GB:Z96072; GB:AL123456; NID:g3261793; PIDN:CAB09489.1; PID:g2

A:Experimental source: strain H37Rv

C:Genetics:

A:Gene: Rv2699c

C:Superfamily: Mycobacterium tuberculosis hypothetical protein Rv2699c

Query Match 100.0%; Score 23; DB 2; Length 100;

Best Local Similarity 100.0%; Pred. No. 1.7e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
||||

Db 4 DYDA 7

RESULT 16

D87037

conserved hypothetical protein ML1026 [imported] - Mycobacterium leprae

C:Species: Mycobacterium leprae

C>Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 10-May-2001

C:Accession: D87037

R:Coile, S.T.; Eiglmeier, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.
R.; Davies, R.M.; Devlin, K.; Duthoy, S.; Feltwell, T.; Fraser, A.; Hamlin, N.;
eam, M.A.; Rutherford, K.M.

Nature 409, 1007-1011, 2001

A:Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares,

A:Title: Massive gene decay in the leprosy bacillus.

A:Reference number: A86909; MUID:21128732; PMID:11234002

A:Accession: D87037

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-100 <STO>

A:Cross-references: GB:AL450380; NID:g13093046; PIDN:CAC31407.1; GSPDB:GN00147

C:Genetics:

A:Gene: ML1026

C:Superfamily: Mycobacterium tuberculosis hypothetical protein Rv2699c

Query Match 100.0%; Score 23; DB 2; Length 100;

Best Local Similarity 100.0%; Pred. No. 1.7e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
||||

Db 4 DYDA 7

RESULT 17

T27821

hypothetical protein ZK287.3 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 04-Mar-2000

C:Accession: T27821

R:McMurray, A.

submitted to the EMBL Data Library, April 1996

A:Reference number: Z20425

A:Accession: T27821

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-102 <WIL>

A:Cross-references: EMBL:Z70757; PIDN:CAA94799.1; GSPDB:GN00023; CESP:ZK287.3

A:Experimental source: clone ZK287

C:Genetics:

A:Gene: CESP:ZK287.3
A:Map position: 5
A:Introns: 57/2
C:Superfamily: Caenorhabditis elegans hypothetical protein ZK287.3

Query Match 100.0%; Score 23; DB 2; Length 102;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4

||||

DB 52 DYDA 55

RESULT 18

WIBP22
gene 17 protein - phage P22

C:Species: phage P22
C:Date: 31-Mar-1991 #sequence_revision 30-Sep-1991 #text_change 23-Jul-1999

C:Accession: S04246

R:Genetician: A.V.; Malloy, D.C.; Poteete, A.R.

J. Mol. Biol. 207, 1-13, 1989

A:Title: Genetic structure of the bacteriophage P22 P(L) operon.

A:Reference number: S04245; MUID:89293845; PMID:2738922

A:Accession: S04246

A:Molecule type: DNA

A:Residues: 1-103 <SEM>

A:Cross-references: EMBL:X15637; NID:g15646; PIDN:CAA33650.1; PID:g15650

A:Note: the authors translated the codon GGT for residue 26 as Gln

C:Genetics:

A:Gene: 17

C:Superfamily: phage P22 gene 17 protein

Query Match

100.0%; Score 23; DB 1; Length 103;

Best Local Similarity 100.0%; Pred. No. 1.7e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4

||||

DB 35 DYDA 38

RESULT 19

S24187

aspartic proteinase inhibitor - potato (fragment)

C:Species: Solanum tuberosum (potato)

C:Date: 19-Mar-1997 #sequence_revision 10-Oct-1997 #text_change 05-Dec-1998

C:Accession: S24187

R:Strukelj, B.; Pungercar, J.; Mesko, P.; Barlic-Maganja, D.; Gubensek, F.; Kregar, I.;

Biol. Chem. Hoppe-Seyler 373, 477-482, 1992

A:Title: Characterization of aspartic proteinase inhibitors from potato at the gene, cDN

A:Reference number: S23961; MUID:92384955; PMID:1515078

A:Accession: S24187

A>Status: not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-103 <STR>

C:Superfamily: cathepsin D inhibitor

C:Keywords: aspartic proteinase inhibitor

F,64-73/Disulfide bonds: #status predicted

Query Match

100.0%; Score 23; DB 2; Length 103;

Best Local Similarity 100.0%; Pred. No. 1.7e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4

||||

DB 17 DYDA 20

RESULT 20

AF1985

nitrogen fixation protein [imported] - Nostoc sp. (strain PCC 7120)

C:Species: Nostoc sp.

A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 30-Jun-2002
C:Accession: AF1985

R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; I;
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Ta
DNA Res. 8, 205-213, 2001

A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacter

A:Reference number: AB1807; MUID:21595285; PMID:11759840

A:Accession: AF1985

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-105 <KUR>

A:Cross-references: GB:BA000019; PIDN:BA073390.1; PID:g17130780; GSPDB:GN00179

A:Experimental source: strain PCC 7120

C:Genetics:

A:Gene: nifW

C:Superfamily: nitrogenase nifW chain

Query Match 100.0%; Score 23; DB 2; Length 105;

Best Local Similarity 100.0%; Pred. No. 1.7e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4

||||

DB 24 DYDA 27

RESULT 21

G69903

hypothetical protein yodL - Bacillus subtilis

C:Species: Bacillus subtilis

C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 28-Jul-2000

C:Accession: G69903

R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.;

C.; Brogl, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N

A.; Ertllich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Feira

Nucleic 390, 249-256, 1997

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.;

Lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hull

Kostter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lar

Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Por

Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; S

A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Segiguchi, J.; Sekowska, U

akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; U

T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yos

A:Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.

A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus sub

A:Reference number: A69580; MUID:98044033; PMID:9384377

A:Accession: G69903

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-105 <KUN>

A:Cross-references: GB:299114; GB:AL009126; NID:g2634230; PIDN:CAB13855.1; PID:el

A:Experimental source: strain 168

C:Genetics:

A:Gene: yodL

C:Superfamily: Bacillus subtilis hypothetical protein yodL

Query Match

100.0%; Score 23; DB 2; Length 105;

Best Local Similarity 100.0%; Pred. No. 1.7e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4

||||

DB 63 DYDA 66

RESULT 22

AG3321

hypothetical protein BMEI0557 [imported] - Brucella melitensis (strain 16M)

C:Species: Brucella melitensis

C;Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 01-Feb-2002
C;Accession: AG3321
R;DelVecchio, V.G.; Kapratral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,
.; Mazur, M.; Goltzman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letess
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A;Title: The genome sequence of the facultative intracellular pathogen *Brucella melitens*
A;Reference number: AD3252; PMID:11756688
A;Accession: AG3321
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-107 <KUR>
A;Cross-references: GB:AE008917; PIDN: AAL51738.1; PID: g17982475; GSPDB: GN00190
A;Experimental source: strain 16M
C;Genetics:
A;Gene: BME10557
A;Map position: I

Query Match 100.0%; Score 23; DB 2; Length 107;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
|||||
DB 62 DYDA 65

RESULT 23
T01654
ribosomal protein l23 - maize (fragment)
N;Alternate names: ribosomal protein l25
C;Species: Zea mays (maize)
C;Date: 19-Feb-1999 #sequence_revision 19-Feb-1999 #text_change 13-Aug-1999
C;Accession: T01654
R;Finkelstein, D.B.; Drew, M.C.; Wing, R.A.; Mullet, J.E.; Jordan, W.R.; Morgan, P.W.
submitted to the EMBL Data Library, April 1998
A;Reference number: Z14371
A;Accession: T01654
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-109 <FIN>
A;Cross-references: EMBL:AF061508; NID: g3264604; PIDN: AAC24573.1; PID: g3264605
A;Experimental source: strain TX5585
C;Superfamily: rat ribosomal protein l23a
C;Keywords: protein biosynthesis; ribosome; RNA binding

Query Match 100.0%; Score 23; DB 2; Length 109;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
|||||
DB 96 DYDA 99

RESULT 24
H87435
transcription regulator, Arsr family [imported] - *Caulobacter crescentus*
C;Species: *Caulobacter crescentus*
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001
C;Accession: H87435
R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.F.; Deboy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A;Title: Complete Genome Sequence of *Caulobacter crescentus*.
A;Reference number: A87249; MUID: 21173698; PMID: 11259647
A;Accession: H87435
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-109 <STO>
A;Cross-references: GB:AE005673; NID: g13422882; PIDN: AAK23484.1; GSPDB: GN00148
C;Genetics:
A;Gene: CC1505

Query Match 100.0%; Score 23; DB 2; Length 109;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
|||||
DB 68 DYDA 71

RESULT 25
C84357
hypothetical protein Vng2074h [imported] - *Halobacterium* sp. NRC-1
C;Species: *Halobacterium* sp. NRC-1
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C;Accession: C84357
R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; La
; Leitthausser, B.; Keller, K.; Cruz, R.; Danison, M.J.; Hough, D.W.; Maddocks, D.G.,
Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T
A;Title: Genome sequence of *Halobacterium* species NRC-1.
A;Reference number: A84160; MUID: 20504483; PMID: 11016950
A;Accession: C84357
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-109 <STO>
A;Cross-references: GB:AE004437; NID: g10581492; PIDN: AAG20223.1; GSPDB: GN00138
C;Genetics:
A;Gene: VNG2074H

Query Match 100.0%; Score 23; DB 2; Length 109;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
|||||
DB 18 DYDA 21

RESULT 26
S19976
Ig kappa chain V region (M-T413) - mouse (fragment)
C;Species: *Mus musculus* (house mouse)
C;Date: 06-Feb-1995 #sequence_revision 06-Feb-1995 #text_change 21-Jan-2000
C;Accession: S19976
R;Weissenhorn, W.; Riettmueller, G.; Weiss, E.M.; Rieber, E.P.
submitted to the EMBL Data Library, March 1992
A;Description: Structural characterization of CD4 mAb.
A;Reference number: S19963
A;Accession: S19976
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-112 <WEI>
A;Cross-references: EMBL: X65093; NID: g52298; PIDN: CAA46221.1; PID: g52299
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: heterotetramer; immunoglobulin
F;16-94/Domain: immunoglobulin homology <IMM>

Query Match 100.0%; Score 23; DB 2; Length 112;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
|||||
DB 30 DYDA 33

RESULT 27
T12731
hypothetical protein 15 - *Methanobacterium* phage psiM2
C;Species: *Methanobacterium* phage psiM2
C;Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 05-May-2000

C:Accession: T12731
R:Pfister, P.; Wasserfallen, A.; Stettler, R.; Leisinger, T.
submitted to the EMBL Data Library, May 1998
A:Description: Archaeophage PsiM2 complete genomic DNA.
A:Reference number: 217578

A:Accession: T12731
A:Status: translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-112 <PFI>

A:Cross-references: EMBL:AF065411; NID:g3249585; PID:g3249600; PIDN:AAC27054.1
A:Experimental source: host Methanobacterium thermoautotrophicum strain Marburg
C:Superfamily: Methanobacterium phage PsiM2 hypothetical protein 15

Query Match 100.0%; Score 23; DB 2; Length 112;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4

Db 65 DYDA 68

RESULT 28

F87193
probable secreted protein [imported] - Mycobacterium leprae

C:Species: Mycobacterium leprae

C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001

C:Accession: F87193

R: Cole, S.T.; Eiglmeier, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; Holt,
R.; Davies, R.M.; Devlin, K.; Duthoy, S.; Feltwell, T.; Fraser, A.; Hamlin, N.; Holroyd,
eam, M.A.; Rutherford, K.M.
Nature 409, 1007-1011, 2001

A:Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.; Squares, S.

A:Title: Massive gene decay in the leprosy bacillus.

A:Reference number: A86909; MUID:21128732; PMID:11234002

A:Accession: F87193

A:Status: Preliminary

A:Molecule type: DNA

A:Residues: 1-112 <STO>

A:Cross-references: GB:AL450380; NID:gl3093911; PIDN:CAC31790.1; GSPDB:GN00147

C:Genetics:
A:Gene: ML2274

Query Match 100.0%; Score 23; DB 2; Length 112;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4

Db 27 DYDA 30

RESULT 29

B70549

hypothetical protein Rv0559c - Mycobacterium tuberculosis (strain H37RV)

C:Species: Mycobacterium tuberculosis

C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999

C:Accession: B70549

R: Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.;
Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.;
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998

A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome

A:Reference number: A70500; MUID:98295987; PMID:9634230

A:Accession: B70549

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-112 <COL>

A:Cross-references: GB:Z95558; GB:AL123456; NID:g3261781; PIDN:CAB08970.1; PID:e1299976;

A:Experimental source: strain H37RV

C:Genetics:

A:Gene: Rv0559c

Query Match 100.0%; Score 23; DB 2; Length 112;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4

Db 27 DYDA 30

RESULT 30

D87301

conserved hypothetical protein CC0421 [imported] - Caulobacter crescentus

C:Species: Caulobacter crescentus

C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 10-May-2001

C:Accession: D87301

R: Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg,
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.;
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser,
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A:Title: Complete Genome Sequence of Caulobacter crescentus.

A:Reference number: A87249; MUID:21173698; PMID:11259647

A:Accession: D87301

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-114 <STO>

A:Cross-references: GB:AE005673; NID:gl3421584; PIDN:AAK22408.1; GSPDB:GN00148

C:Genetics:

A:Gene: CC0421

C:Superfamily: hypothetical protein HI0925

Query Match 100.0%; Score 23; DB 2; Length 114;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4

Db 69 DYDA 72

RESULT 31

JE0178

SH3 binding glutamate-rich protein - human

C:Species: Homo sapiens (man)

C:Date: 03-Jul-1998 #sequence_revision 10-Jul-1998 #text_change 21-Jul-2000

C:Accession: JE0178

R: Egeo, A.; Mazzocco, M.; Arrigo, P.; Vidal-Taboada, J.M.; Oliva, R.; Pirola, B.;
Biochem. Biophys. Res. Commun. 247, 302-306, 1998

A:Title: Identification and characterization of a new human gene encoding a small
A:Reference number: JE0178; MUID:98308110; PMID:9642120

A:Accession: JE0178

A:Molecule type: mRNA

A:Residues: 1-114 <EGE>

A:Cross-references: GB:AF042081; NID:g3337419; PIDN:AAC27445.1; PID:g3337420

C:Genetics:

A:Gene: SH3BGR

A:Map position: Xq13.3

C:Superfamily: human SH3 binding glutamate-rich protein

F:59-67/Region: proline-rich SH3 binding motif

Query Match 100.0%; Score 23; DB 2; Length 114;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4

Db 78 DYDA 81

RESULT 32

D30540

Ig heavy chain V region (174.3F4) - mouse (fragment)

C:Species: Mus musculus (house mouse)

C;Date: 10-Feb-1989 #sequence_revision 10-Feb-1989 #text_change 16-Aug-1996

C;Accession: D30540

R;Claflin, J.L.; Berry, J.

J. Immunol. 141, 4012-4019, 1988

A;Title: Genetics of the phosphocholine-specific antibody response to Streptococcus pneumoniae

A;Reference number: A30534; PMID:8903545; PMID:3141511

A;Accession: D30540

A;Status: preliminary; nucleic acid sequence not shown; not compared with conceptual translation

A;Molecule type: mRNA

A;Residues: 1-119 <CL>

C;Superfamily: immunoglobulin V region; immunoglobulin homology

C;Keywords: heterotetramer; immunoglobulin

F;15-100/Domain: immunoglobulin homology <IMM>

Query Match 100.0%; Score 23; DB 2; Length 119;

Best Local Similarity 100.0%; Pred. No. 2e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DYDA 4

Db 102 DYDA 105

RESULT 33

AB1425

hypothetical protein lmo2803 [imported] - Listeria monocytogenes (strain EGD-e)

C;Species: Listeria monocytogenes

C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001

C;Accession: AB1425

R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihl, H.

D.; Jones, L.M.; Karat, U.

Science 294, 849-852, 2001

A;Authors: Kref, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Maok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A.

A;Title: Comparative genomics of Listeria species

A;Reference number: AB1077; PMID:21537279; PMID:11679669

A;Accession: AB1425

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-120 <GLA>

A;Cross-references: GB:NC_003210; PIDN:CAD01016.1; PID:gl6412303; GSPDB:GN00177

C;Experimental source: strain EGD-e

C;Genetics:

A;Gene: lmo2803

Query Match

Best Local Similarity 100.0%; Score 23; DB 2; Length 120;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DYDA 4

Db 58 DYDA 61

RESULT 34

G96983

nitrogen regulatory protein PII, gene nrgB [imported] - Clostridium acetobutylicum

C;Species: Clostridium acetobutylicum

C;Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 30-Sep-2001

C;Accession: G96983

R;Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee, J.

J. Bacteriol. 183, 4823-4838, 2001

A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clostridium acetobutylicum

A;Reference number: A96900; PMID:21359325; PMID:21359325

A;Accession: G96983

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-121 <KUR>

A;Cross-references: GB:AE001437; PIDN:AAK78658.1; PID:gl5023558; GSPDB:GN00168

A;Experimental source: Clostridium acetobutylicum ATCC824

C;Genetics:

A;Gene: CAC0681

C;Superfamily: regulatory protein P-II

Query Match 100.0%; Score 23; DB 2; Length 121;

Best Local Similarity 100.0%; Pred. No. 2e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DYDA 4

Db 112 DYDA 115

RESULT 35

F87429

hypothetical protein CC1453 [imported] - Caulobacter crescentus

C;Species: Caulobacter crescentus

C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001

C;Accession: F87429

R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.;

n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A;Title: Complete Genome Sequence of Caulobacter crescentus.

A;Reference number: A87249; PMID:21173698; PMID:11259647

A;Accession: F87429

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-124 <STO>

A;Cross-references: GB:AE005673; NID:gl3422820; PIDN:AAK23434.1; GSPDB:GN00148

C;Genetics:

A;Gene: CC1453

Query Match

Best Local Similarity 100.0%; Score 23; DB 2; Length 124;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DYDA 4

Db 77 DYDA 80

RESULT 36

F84355

hypothetical protein Vng2054h [imported] - Halobacterium sp. NRC-1

C;Species: Halobacterium sp. NRC-1

C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001

C;Accession: F84355

R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; La

; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.;

Jung, K.H.; Alam, M.; Freitas, T.

Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000

A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.

A;Title: Genome sequence of Halobacterium species NRC-1.

A;Reference number: A84160; PMID:20504483; PMID:11016950

A;Accession: F84355

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-127 <STO>

A;Cross-references: GB:AE004437; NID:gl0581477; PIDN:AAG20210.1; GSPDB:GN00138

C;Genetics:

A;Gene: VNG2054H

Query Match

Best Local Similarity 100.0%; Score 23; DB 2; Length 127;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DYDA 4

Db 31 DYDA 34

RESULT 37

G87615

conserved hypothetical protein CC2961 [imported] - Caulobacter crescentus
 C:Species: Caulobacter crescentus
 C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001
 C:Accession: G87615

R.Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
 n, J.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
 Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A:Title: Complete Genome Sequence of Caulobacter crescentus.

A:Reference number: A87249; MUID:21173698; PMID:11259647

A:Accession: G87615

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-129 <STO>

A:Cross-references: GB:AE005673; NID:g13424591; PIDN:AAK24923.1; GSPDB:GN00148

C:Genetics:

A:Gene: CC2961

Query Match 100.0%; Score 23; DB 2; Length 129;

Best Local Similarity 100.0%; Pred. No. 2.2e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4

||||

Db 23 DYDA 26

RESULT 38

E97875

degenerate transposase (orf1) [imported] - Streptococcus pneumoniae (strain R6)

C:Species: Streptococcus pneumoniae

C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 22-Oct-2001

C:Accession: E97875

R.Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; H

e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M

y, P.; Sun, P.M.; Winkler, M.E.

J. Bacteriol. 183, 5709-5717, 2001

A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;

A:Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.

A:Reference number: A97872; MUID:21429245; PMID:11544234

A:Accession: E97875

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-130 <KUR>

A:Cross-references: GB:AE007317; PIDN:AAK98833.1; PID:g15457559; GSPDB:GN00174

C:Genetics:

A:Gene: IS1167-truncation

Query Match 100.0%; Score 23; DB 2; Length 130;

Best Local Similarity 100.0%; Pred. No. 2.2e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4

||||

Db 37 DYDA 40

RESULT 39

VCBPPI

coat protein - phage PRRI

C:Species: phage PRRI

C:Date: 30-Nov-1979 #sequence_revision 30-Nov-1979 #text_change 31-Dec-1993

C:Accession: A04225

R.Dhaese, P.; Vandekerckhove, J.S.; Van Montagu, M.C.

Eur. J. Biochem. 94, 375-386, 1979

A:Title: The primary structure of the coat protein of the broad-host-range RNA bacterioph

A:Reference number: A04225; MUID:79148387; PMID:107028

A:Accession: A04225

A:Molecule type: protein

A:Residues: 1-131 <DHA>

C:Superfamily: phage GA coat protein

Query Match 100.0%; Score 23; DB 1; Length 131;

Best Local Similarity 100.0%; Pred. No. 2.2e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4

||||

Db 92 DYDA 95

RESULT 40

AG2739

conserved hypothetical protein Atul327 [imported] - Agrobacterium tumefaciens (str

C:Species: Agrobacterium tumefaciens

C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 11-Jan-2002

C:Accession: AG2739

R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.;

erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; M

; Karp, P.; Romero, P.; Zhang, S.

Science 294, 2317-2323, 2001

A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon

ster, E.W.

A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58

A:Reference number: AB2577; PMID:11743193

A:Accession: AG2739

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-132 <KUR>

A:Cross-references: GB:AE008688; PIDN:AAK42333.1; PID:g17739737; GSPDB:GN00186

A:Experimental source: strain C58 (Dupont)

C:Genetics:

A:Gene: Atul327

A:Map position: circular chromosome

Query Match 100.0%; Score 23; DB 2; Length 132;

Best Local Similarity 100.0%; Pred. No. 2.2e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4

||||

Db 124 DYDA 127

RESULT 41

E97520

13kDa differentiation-associated protein [imported] - Agrobacterium tumefaciens (s

C:Species: Agrobacterium tumefaciens

C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 11-Jan-2002

C:Accession: E97520

R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorollo, B.; G

A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Marke

Science 294, 2323-2328, 2001

A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacter

A:Reference number: A97359; PMID:11743194

A:Accession: E97520

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-132 <KUR>

A:Cross-references: GB:AE007869; PIDN:AAK87118.1; PID:g15156382; GSPDB:GN00169

C:Genetics:

A:Gene: AGR_C_2444

A:Map position: circular chromosome

Query Match 100.0%; Score 23; DB 2; Length 132;

Best Local Similarity 100.0%; Pred. No. 2.2e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4

||||

Db 124 DYDA 127

RESULT 42

F70659

hypothetical protein Rv2546 - Mycobacterium tuberculosis (strain H37RV)

C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
C:Accession: F70659
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nature 393, 537-544, 1998
A:Authors: Sqaers, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: F70659
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-137 <COL>
A:Cross-references: GB:283863; GB:ALJ23456; NID:g3261685; PIDN:CAB06192.1; PID:e290871;
A:Experimental source: strain H37Rv
C:Genetics:
A:Gene: RV2546

Query Match 100.0%; Score 23; DB 2; Length 137;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DYDA 4
|||||
Db 50 DYDA 53

RESULT 43
T04904
hypothetical protein T10114.60 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-Jun-1999
C:Accession: T04904
R:Bevan, M.; Murphy, G.; Ridley, P.; Hudson, S.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X submitted to the Protein Sequence Database, April 1998
A:Reference number: Z15389
A:Accession: T04904
A:Molecule type: DNA
A:Residues: 1-140 <BEV>
A:Cross-references: EMBL:AL021712
A:Experimental source: cultivar Columbia; BAC clone T10114
C:Genetics:
A:Map position: 4
A:Introns: 26/1; 73/1
A:Note: T10114.60

Query Match 100.0%; Score 23; DB 2; Length 140;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DYDA 4
|||||
Db 127 DYDA 130

RESULT 44
RBY25
ribosomal protein L23a.e, cytosolic - yeast (Saccharomyces cerevisiae)
N:Alternate names: protein AOX142i; protein O0534; protein YOL127w; ribosomal protein rp
C:Species: Saccharomyces cerevisiae
C:Date: 17-Mar-1987 #sequence_revision 19-Jul-1996 #text_change 16-Jun-2000
C:Accession: S63443; A02820; S45520; S66824; S71981
R:Casamayor, A.; Khalid, H.; Balcells, L.; Aldea, M.; Casas, C.; Herrero, E.; Arino, J. submitted to the EMBL Data Library, November 1995
A:Description: Sequencing of a 13.4 kbp fragment of the left arm of chromosome XV reveal
A:Reference number: S63440
A:Accession: S63443
A:Molecule type: DNA
A:Residues: 1-142 <CAS>
A:Cross-references: EMBL:U41293; NID:gl209710; PIDN:AAC49465.1; PID:gl209714
A:Experimental source: strain FX1679
R:Leer, R.J.; van Raamsdonk-Duin, M.M.C.; Hagendoorn, M.J.M.; Mager, W.H.; Planta, R.J.

Nucleic Acids Res. 12, 6685-6700, 1984

A:Title: Structural comparison of yeast ribosomal protein genes.
A:Reference number: A93332; MUID:65014125; PMID:6091033
A:Accession: A02820
A:Molecule type: DNA
A:Residues: 1-105, 'NI', 113-142 <LEE>
A:Cross-references: EMBL:X01014; NID:g4325; PIDN:CAA25506.1; PID:g4326
R:Takahura, H.; Tsunasawa, S.; Miyagi, M.; Warner, J.R. J. Biol. Chem. 267, 5443-5445, 1992
A:Title: NH2-terminal acylation of ribosomal proteins of Saccharomyces cerevisia
A:Reference number: S45500; MUID:92184799; PMID:1544921
A:Accession: S45520
A:Molecule type: protein
A:Residues: 2-21 <TAK>
R:Arino, J.; Casamayor, A.; Gamo, F.J.; Gancedo, C.; Lafuente, M.J.; Aldea, M.; Ca submitted to the Protein Sequence Database, July 1996
A:Reference number: S66814
A:Accession: S66824
A:Molecule type: DNA
A:Residues: 1-142 <ARI>
A:Cross-references: EMBL:Z74869; GSPDB:GNO0015; MIPS:YOL127w; NID:gl420010; PIDN:C
A:Experimental source: strain S288C
R:Casamayor, A.; Khalid, H.; Balcells, L.; Aldea, M.; Casas, C.; Herrero, E.; Arin Yeast 12, 1013-1020, 1996
A:Title: Sequence analysis of a 13.4 kbp fragment from the left arm of chromosome eading frames.
A:Reference number: S71978; MUID:97051588; PMID:8896265
A:Accession: S71981
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-142 <CAW>
A:Cross-references: EMBL:U41293; NID:gl209710; PIDN:AAC49465.1; PID:gl209714
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1 C:Genetics:
A:Gene: SGD:RPL25; YL25; MIPS:YOL127w
A:Cross-references: SGD:S0005487; MIPS:YOL127w
A:Map position: 15L
A:Introns: 5/1
A:Superfamily: rat ribosomal protein L23a
C:Keywords: protein biosynthesis; ribosome; RNA binding
F:2-142/Product: ribosomal protein L23a.e #status experimental <MAT>

Query Match 100.0%; Score 23; DB 1; Length 142;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DYDA 4
|||||
Db 129 DYDA 132

RESULT 45
S30000
ribosomal protein L23a.e - yeast (Kluyveromyces marxianus)
N:Alternate names: ribosomal protein YL25
C:Species: Kluyveromyces marxianus
C:Date: 07-May-1993 #sequence_revision 07-May-1993 #text_change 13-Aug-1999
C:Accession: S30000
R:Bergkamp-Steffens, G.K.; Hoekstra, R.; Planta, R.J. yeast 8, 903-922, 1992
A:Title: Structural and putative regulatory sequences of Kluyveromyces ribosomal p
A:Reference number: S29999; MUID:93127726; PMID:1481569
A:Accession: S30000
A:Molecule type: DNA
A:Residues: 1-142 <BER>
A:Cross-references: GB:S53422; NID:g263476; PIDN:AAB24897.1; PID:g263477
A:Note: the authors did not show the third nucleotide in the codon for 10-Ala
C:Genetics:
A:Gene: L25
A:Introns: 5/1
A:Superfamily: rat ribosomal protein L23a
C:Keywords: protein biosynthesis; ribosome; RNA binding

Query Match 100.0%; Score 23; DB 2; Length 142;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
Db 129 DYDA 132

Search completed: February 6, 2003, 11:23:02
Job time : 14.5 secs

GenCore version 5.1.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 6, 2003, 11:19:49 ; Search time 14 Seconds
(without alignments)
23.165 Million cell updates/sec

Title: PAT943-1

Perfect score: 23

Sequence: 1 dyda 4

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 405691 seqs, 81078759 residues

Total number of hits satisfying chosen parameters: 483

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 100%

Maximum Match 100%

Listing first 45 summaries

Database :

- Pending_Patents_AA_New.*
- 1: /cgn2_6/ptodata/2/paa/PCT_NEW_COMB.pcp.*
 - 2: /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pcp.*
 - 3: /cgn2_6/ptodata/2/paa/US07_NEW_COMB.pcp.*
 - 4: /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pcp.*
 - 5: /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pcp.*
 - 6: /cgn2_6/ptodata/2/paa/US10_NEW_COMB.pcp.*
 - 7: /cgn2_6/ptodata/2/paa/US60_NEW_COMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	23	100.0	52	6	US-10-218-102-323
2	23	100.0	58	6	US-10-276-774-1476
3	23	100.0	61	1	PCT-US02-32727-28253
4	23	100.0	61	6	US-10-057-498-28253
5	23	100.0	79	6	US-10-092-411A-3420
6	23	100.0	92	1	PCT-US02-32727-27674
7	23	100.0	92	6	US-10-057-498-27674
8	23	100.0	94	5	US-09-513-999C-4805
9	23	100.0	95	6	US-10-218-140-1452
10	23	100.0	98	5	US-09-513-999C-5067
11	23	100.0	99	1	PCT-US02-32727-19310
12	23	100.0	99	6	US-10-057-498-19310
13	23	100.0	102	5	US-09-950-084-6433
14	23	100.0	108	1	PCT-US02-32727-21054
15	23	100.0	108	6	US-10-057-498-21054
16	23	100.0	114	5	US-09-724-676-51371
17	23	100.0	114	5	US-09-724-676A-51371
18	23	100.0	114	6	US-10-319-763-214
19	23	100.0	115	6	US-10-319-763-120
20	23	100.0	119	1	PCT-US02-22011-121
21	23	100.0	119	1	PCT-US02-22011-122
22	23	100.0	119	6	US-10-194-975-121
23	23	100.0	119	6	US-10-194-975-122
24	23	100.0	128	6	US-10-264-237-1823
25	23	100.0	130	1	PCT-US02-32727-5855
26	23	100.0	130	6	US-10-057-498-5855

27	23	100.0	145	5	US-09-724-676-51020	Sequence 51020, A
28	23	100.0	145	5	US-09-724-676A-51020	Sequence 51020, A
29	23	100.0	159	5	US-09-950-084-4003	Sequence 4003, Ap
30	23	100.0	169	6	US-10-264-237-1916	Sequence 1916, Ap
31	23	100.0	170	1	PCT-US02-32727-23164	Sequence 23164, A
32	23	100.0	170	6	US-10-057-498-23164	Sequence 23164, A
33	23	100.0	171	6	US-10-276-774-2363	Sequence 2363, Ap
34	23	100.0	173	5	US-09-950-084-5911	Sequence 5911, Ap
35	23	100.0	175	5	US-09-724-676-51015	Sequence 51015, A
36	23	100.0	175	5	US-09-724-676A-51015	Sequence 51015, A
37	23	100.0	175	5	US-09-950-084-3836	Sequence 3836, Ap
38	23	100.0	202	5	US-09-134-000C-4718	Sequence 4718, Ap
39	23	100.0	202	5	US-09-134-000C-4718	Sequence 4718, Ap
40	23	100.0	206	5	US-09-134-000C-3471	Sequence 3471, Ap
41	23	100.0	206	5	US-09-134-000C-3471	Sequence 3471, Ap
42	23	100.0	225	5	US-09-724-676-67070	Sequence 67070, A
43	23	100.0	225	5	US-09-724-676-67071	Sequence 67071, A
44	23	100.0	225	5	US-09-724-676-67072	Sequence 67072, A
45	23	100.0	225	5	US-09-724-676-67073	Sequence 67073, A

ALIGNMENTS

RESULT 1
US-10-218-102-323
; Sequence 323, Application US/10218102
; GENERAL INFORMATION:
; APPLICANT: Bentzien, Joerg
; APPLICANT: Dahiyat, Bassil I.
; APPLICANT: Desjarlais, John R.
; APPLICANT: Hayes, Robert J.
; APPLICANT: Viemetter, Jost
; TITLE OF INVENTION: Protein Design Automation for Protein Libraries
; FILE REFERENCE: A-67229-11/RFT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/10/218.102
; CURRENT FILING DATE: 2002-08-12
; PRIOR APPLICATION NUMBER: US 09/927,790
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: US 60/311,545
; PRIOR FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: US 60/324,899
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 60/351,937
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: US 60/352,103
; PRIOR FILING DATE: 2002-01-25
; SOFTWARE: Patent version 3.1
; SEQ ID NO 323
; LENGTH: 52
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: synthetic
US-10-218-102-323

Query Match 100.0%; Score 23; DB 6; Length 52;
Best Local Similarity 100.0%; Pred. No. 99;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 DYDA 4
Db 5 DYDA 8

RESULT 2

US-10-276-774-1476
; Sequence 1476, Application US/10276774
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; APPLICANT: Tang, Y, Tom et al
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides

```
; FILE REFERENCE: 21272-030
; CURRENT APPLICATION NUMBER: US/10/276,774
; CURRENT FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 2700
; SOFTWARE: Custom
; SEQ ID NO 1476
; LENGTH: 58
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(58)
; OTHER INFORMATION: Xaa = any amino acid or nothing
US-10-276-774-1476

Query Match          100.0%; Score 23; DB 6; Length 58;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
Db 20 DYDA 23

RESULT 3
PCT-US02-32727-28253
; Sequence 28253, Application PC/TUS0232727
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer
; APPLICANT: Skeiky, Yasir
; APPLICANT: Persing, David
; APPLICANT: Bhatia, Ajay
; APPLICANT: Maisonneuve, Jean Francois
; APPLICANT: Zhang, Yanni
; APPLICANT: Wang, Siqing
; APPLICANT: Jen, Shyian
; APPLICANT: Lodes, Michael
; APPLICANT: Benson, Darin
; APPLICANT: Jones, Robert
; APPLICANT: Carter, Darrick
; APPLICANT: Barth, Brenda
; APPLICANT: Douglass, John
; TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acne
; FILE REFERENCE: 210121.514C1
; CURRENT APPLICATION NUMBER: PCT/US02/32727
; CURRENT FILING DATE: 2002-10-11
; NUMBER OF SEQ ID NOS: 30992
; SEQ ID NO 28253
; LENGTH: 61
; TYPE: PRT
; ORGANISM: Propioni acnes
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (2)
; OTHER INFORMATION: Xaa = Any Amino Acid
PCT-US02-32727-28253
```

```
Query Match          100.0%; Score 23; DB 1; Length 61;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 DYDA 4
Db 46 DYDA 49
```

```
RESULT 4
US-10-057-498-28253
; Sequence 28253, Application US/10057498
```

```
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer
; APPLICANT: Skeiky, Yasir
; APPLICANT: Persing, David
; TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of
; FILE REFERENCE: 210121.514
; CURRENT APPLICATION NUMBER: US/10/057,498
; CURRENT FILING DATE: 2001-04-20
; NUMBER OF SEQ ID NOS: 29212
; SEQ ID NO 28253
; LENGTH: 61
; TYPE: PRT
; ORGANISM: Propioni acnes
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (2)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-10-057-498-28253
```

```
Query Match          100.0%; Score 23; DB 6; Length 61;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 DYDA 4
Db 46 DYDA 49
```

```
RESULT 5
US-10-092-411A-3420
; Sequence 3420, Application US/10092411A
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLO
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032798-101
; CURRENT APPLICATION NUMBER: US/10/092,411A
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: US 09/134,001
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5676
; SEQ ID NO 3420
; LENGTH: 79
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-10-092-411A-3420
```

```
Query Match          100.0%; Score 23; DB 6; Length 79;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 DYDA 4
Db 47 DYDA 50
```

```
RESULT 6
PCT-US02-32727-27674
; Sequence 27674, Application PC/TUS0232727
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer
; APPLICANT: Skeiky, Yasir
; APPLICANT: Persing, David
; APPLICANT: Bhatia, Ajay
; APPLICANT: Maisonneuve, Jean Francois
; APPLICANT: Zhang, Yanni
; APPLICANT: Wang, Siqing
; APPLICANT: Jen, Shyian
; APPLICANT: Lodes, Michael
```

; APPLICANT: Benson, Darin
; APPLICANT: Jones, Robert
; APPLICANT: Carter, Darrick
; APPLICANT: Barth, Brenda
; APPLICANT: Douglass, John
; TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acne
; FILE REFERENCE: 210121.514C1
; CURRENT APPLICATION NUMBER: PCT/US02/32727
; CURRENT FILING DATE: 2002-10-11
; NUMBER OF SEQ ID NOS: 30992
; SEQ ID NO 27674
; LENGTH: 92
; TYPE: PRT
; ORGANISM: Propionibacterium acnes
PCT-US02-32727-27674

Query Match 100.0%; Score 23; DB 1; Length 92;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
|
|
|
|
Db 11 DYDA 14

RESULT 7
US-10-057-498-27674
; Sequence 27674, Application US/10057498
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer
; APPLICANT: Skeiky, Yasir
; APPLICANT: Persing, David
; TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acne
; FILE REFERENCE: 210121.514
; CURRENT APPLICATION NUMBER: US/10/057,498
; CURRENT FILING DATE: 2001-04-20
; NUMBER OF SEQ ID NOS: 29212
; SEQ ID NO 27674
; LENGTH: 92
; TYPE: PRT
; ORGANISM: Propionibacterium acnes
US-10-057-498-27674

Query Match 100.0%; Score 23; DB 6; Length 92;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
|
|
|
|
Db 11 DYDA 14

RESULT 8
US-09-513-999C-4805
; Sequence 4805, Application US/09513999C
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 4805
; LENGTH: 94
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-513-999C-4805

Query Match 100.0%; Score 23; DB 5; Length 94;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
|
|
|
|
Db 71 DYDA 74

RESULT 9
US-10-218-140-1452
; Sequence 1452, Application US/10218140
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; APPLICANT: Shinkets, Richard A.
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES AND POLYPEPTIDES
; TITLE OF INVENTION: ENCODED THEREBY
; FILE REFERENCE: 15966-543 CON
; CURRENT APPLICATION NUMBER: US/10/218,140
; CURRENT FILING DATE: 2002-08-12
; PRIOR APPLICATION NUMBER: 09/540,763
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: 60/127,728
; PRIOR FILING DATE: 1999-04-05
; PRIOR APPLICATION NUMBER: 60/127,636
; PRIOR FILING DATE: 1999-04-02
; PRIOR APPLICATION NUMBER: 60/127,607
; PRIOR FILING DATE: 1999-03-31
; NUMBER OF SEQ ID NOS: 6322
; SOFTWARE: Curanator Version 1.0
; SEQ ID NO 1452
; LENGTH: 95
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-218-140-1452

Query Match 100.0%; Score 23; DB 6; Length 95;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
|
|
|
|
Db 52 DYDA 55

RESULT 10
US-09-513-999C-5067
; Sequence 5067, Application US/09513999C
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 5067
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-513-999C-5067

Query Match 100.0%; Score 23; DB 5; Length 98;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
|
|
|
|
Db 78 DYDA 81

RESULT 11
PCT-US02-32727-19310

; Sequence 19310, Application PC/TUS0232727

; GENERAL INFORMATION:

; APPLICANT: Mitcham, Jennifer

; APPLICANT: Skeiky, Yasir

; APPLICANT: Persing, David

; APPLICANT: Bhatia, Ajay

; APPLICANT: Maisonneuve, Jean Francois

; APPLICANT: Zhang, Yanni

; APPLICANT: Wang, Siqing

; APPLICANT: Jen, Shyian

; APPLICANT: Lodes, Michael

; APPLICANT: Benson, Darin

; APPLICANT: Jones, Robert

; APPLICANT: Carter, Darrick

; APPLICANT: Barth, Brenda

; APPLICANT: Douglass, John

; TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acnes

; FILE REFERENCE: 210121.514C1

; CURRENT APPLICATION NUMBER: PCT/US02/32727

; NUMBER OF SEQ ID NOS: 30992

; SEQ ID NO 19310

; LENGTH: 99

; TYPE: PRT

; ORGANISM: Propioni acnes

PCT-US02-32727-19310

Query Match 100.0%; Score 23; DB 1; Length 99;

Best Local Similarity 100.0%; Pred. No. 1.9e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4

||||

Db 54 DYDA 57

||||

RESULT 12

US-10-057-498-19310

; Sequence 19310, Application US/10057498

; GENERAL INFORMATION:

; APPLICANT: Mitcham, Jennifer

; APPLICANT: Skeiky, Yasir

; APPLICANT: Persing, David

; TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acnes

; FILE REFERENCE: 210121.514

; CURRENT APPLICATION NUMBER: US/10/057,498

; CURRENT FILING DATE: 2001-04-20

; NUMBER OF SEQ ID NOS: 29212

; SEQ ID NO 19310

; LENGTH: 99

; TYPE: PRT

; ORGANISM: Propioni acnes

US-10-057-498-19310

Query Match 100.0%; Score 23; DB 6; Length 99;

Best Local Similarity 100.0%; Pred. No. 1.9e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4

||||

Db 54 DYDA 57

||||

RESULT 13

US-09-950-084-6433

; Sequence 6433, Application US/09950084

; GENERAL INFORMATION:

; APPLICANT: George H. Shimer, Jr.

; APPLICANT: George H. Miller

; APPLICANT: Roberta S. Hare

; APPLICANT: Karen J. Shaw

; TITLE OF INVENTION: Staphylococcus aureus Related Compositions and Methods

; FILE REFERENCE: 1034/IC963US2

; CURRENT APPLICATION NUMBER: US/09/950,084

; CURRENT FILING DATE: 2001-09-10

; PRIOR APPLICATION NUMBER: US 09/417,811

; PRIOR FILING DATE: 1999-10-14

; PRIOR APPLICATION NUMBER: US 09/353,718

; PRIOR FILING DATE: 1999-07-14

; PRIOR APPLICATION NUMBER: US 09/266,557

; PRIOR FILING DATE: 1999-03-11

; PRIOR APPLICATION NUMBER: US 09/266,556

; PRIOR FILING DATE: 1999-03-11

; PRIOR APPLICATION NUMBER: US 09/266,555

; PRIOR FILING DATE: 1999-03-11

; PRIOR APPLICATION NUMBER: US 09/266,542

; PRIOR FILING DATE: 1999-03-11

; PRIOR APPLICATION NUMBER: US 09/266,541

; PRIOR FILING DATE: 1999-03-11

; PRIOR APPLICATION NUMBER: US 09/037,934

; PRIOR FILING DATE: 1998-03-10

; PRIOR APPLICATION NUMBER: US 09/036,720

; PRIOR FILING DATE: 1998-03-06

; PRIOR APPLICATION NUMBER: US 09/036,338

; PRIOR FILING DATE: 1998-03-06

; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 7451

; SEQ ID NO 6433

; LENGTH: 102

; TYPE: PRT

; ORGANISM: Staphylococcus aureus

US-09-950-084-6433

Query Match 100.0%; Score 23; DB 5; Length 102;

Best Local Similarity 100.0%; Pred. No. 2e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4

||||

Db 36 DYDA 39

||||

RESULT 14

PCT-US02-32727-21054

; Sequence 21054, Application PC/TUS0232727

; GENERAL INFORMATION:

; APPLICANT: Mitcham, Jennifer

; APPLICANT: Skeiky, Yasir

; APPLICANT: Persing, David

; APPLICANT: Bhatia, Ajay

; APPLICANT: Maisonneuve, Jean Francois

; APPLICANT: Zhang, Yanni

; APPLICANT: Wang, Siqing

; APPLICANT: Jen, Shyian

; APPLICANT: Lodes, Michael

; APPLICANT: Benson, Darin

; APPLICANT: Jones, Robert

; APPLICANT: Carter, Darrick

; APPLICANT: Barth, Brenda

; APPLICANT: Douglass, John

; TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of A

; FILE REFERENCE: 210121.514C1

; CURRENT APPLICATION NUMBER: PCT/US02/32727

; CURRENT FILING DATE: 2002-10-11

; NUMBER OF SEQ ID NOS: 30992

; SEQ ID NO 21054

; LENGTH: 108

; TYPE: PRT

; ORGANISM: Propioni acnes

; FEATURE:

; NAME/KEY: unsure

; LOCATION: (2)

; OTHER INFORMATION: Xaa = Any Amino Acid
PCT-US02-32727-21054

Query Match 100.0%; Score 23; DB 1; Length 108;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
||||
DB 93 DYDA 96

RESULT 15

US-10-057-498-21054
; Sequence 21054, Application US/10057498

; GENERAL INFORMATION:

; APPLICANT: Mitcham, Jennifer

; APPLICANT: Skeiky, Yasir

; APPLICANT: Persing, David

; TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acne

; FILE REFERENCE: 210121.514

; CURRENT APPLICATION NUMBER: US/10/057,498

; CURRENT FILING DATE: 2001-04-20

; NUMBER OF SEQ ID NOS: 29212

; SEQ ID NO 21054

; LENGTH: 108

; TYPE: PRT

; ORGANISM: Propioni acnes

; FEATURE:

; NAME/KEY: unsure

; LOCATION: (2)

; OTHER INFORMATION: Xaa = Any Amino Acid

US-10-057-498-21054

Query Match

100.0%; Score 23; DB 6; Length 108;

Best Local Similarity 100.0%; Pred. No. 2.1e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
||||
DB 93 DYDA 96

RESULT 16

US-09-724-676-51371

; Sequence 51371, Application US/09724676

; GENERAL INFORMATION:

; APPLICANT: Compugen LTD

; TITLE OF INVENTION: Variants of alternative splicing

; FILE REFERENCE: 129181.4 Compugen

; CURRENT APPLICATION NUMBER: US/09/724,676

; CURRENT FILING DATE: 2000-11-28

; NUMBER OF SEQ ID NOS: 97222

; SOFTWARE: Patentin version 3.2

; SEQ ID NO 51371

; LENGTH: 114

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-724-676-51371

Query Match

100.0%; Score 23; DB 5; Length 114;

Best Local Similarity 100.0%; Pred. No. 2.3e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
||||
DB 78 DYDA 81

RESULT 17

US-09-724-676A-51371

; Sequence 51371, Application US/09724676A

; GENERAL INFORMATION:

; APPLICANT: Compugen LTD

; TITLE OF INVENTION: Variants of alternative splicing

; FILE REFERENCE: 129181.4 Compugen

; CURRENT APPLICATION NUMBER: US/09/724,676A

; CURRENT FILING DATE: 2000-11-28

; NUMBER OF SEQ ID NOS: 97222

; SOFTWARE: Patentin version 3.2

; SEQ ID NO 51371

; LENGTH: 114

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-724-676A-51371

Query Match

100.0%; Score 23; DB 5; Length 114;

Best Local Similarity 100.0%; Pred. No. 2.3e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
||||
DB 78 DYDA 81

RESULT 18

US-10-319-763-214

; Sequence 214, Application US/10319763

; GENERAL INFORMATION:

; APPLICANT: Dumas Milne Edwards, Jean-Baptiste

; APPLICANT: Duclert, Aymeric

; APPLICANT: Bouqueleret, Lydie

; TITLE OF INVENTION: EXTENDED CDNAS FOR SECRETED PROTEINS

; FILE REFERENCE: G-031.US04.DIV

; CURRENT APPLICATION NUMBER: US/10/319,763

; CURRENT FILING DATE: 2002-12-10

; PRIOR APPLICATION NUMBER: 60/066,677

; PRIOR FILING DATE: 1997-11-13

; PRIOR APPLICATION NUMBER: 60/069,957

; PRIOR FILING DATE: 1997-12-17

; PRIOR APPLICATION NUMBER: 60/074,121

; PRIOR FILING DATE: 1998-02-09

; PRIOR APPLICATION NUMBER: 60/081,563

; PRIOR FILING DATE: 1998-04-13

; PRIOR APPLICATION NUMBER: 60/096,116

; PRIOR FILING DATE: 1998-08-10

; PRIOR APPLICATION NUMBER: 60/099,273

; PRIOR FILING DATE: 1998-09-04

; NUMBER OF SEQ ID NOS: 229

; SOFTWARE: Patent.pm

; SEQ ID NO 214

; LENGTH: 114

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: SIGNAL

; LOCATION: -103...-1

US-10-319-763-214

Query Match

100.0%; Score 23; DB 6; Length 114;

Best Local Similarity 100.0%; Pred. No. 2.3e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
||||
DB 78 DYDA 81

RESULT 19

US-10-319-763-120

; Sequence 120, Application US/10319763

; GENERAL INFORMATION:

; APPLICANT: Dumas Milne Edwards, Jean-Baptiste

; APPLICANT: Duclert, Aymeric

; APPLICANT: Bouqueleret, Lydie

; TITLE OF INVENTION: EXTENDED CDNAS FOR SECRETED PROTEINS

; FILE REFERENCE: G-031.US04.DIV
; CURRENT APPLICATION NUMBER: US/10/319,763
; CURRENT FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: 60/066,677
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/069,957
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/074,121
; PRIOR FILING DATE: 1998-02-09
; PRIOR APPLICATION NUMBER: 60/081,563
; PRIOR FILING DATE: 1998-04-13
; PRIOR APPLICATION NUMBER: 60/096,116
; PRIOR FILING DATE: 1998-08-10
; PRIOR APPLICATION NUMBER: 60/099,273
; PRIOR FILING DATE: 1998-09-04
; NUMBER OF SEQ ID NOS: 229
; SOFTWARE: Patent.pm
; SEQ ID NO 120
; LENGTH: 115
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -103...-1
US-10-319-763-120

Query Match 100.0%; Score 23; DB 6; Length 115;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
|
|
|
|
Db 78 DYDA 81

RESULT 20

PCT-US02-22011-121
; Sequence 121, Application PC/TUS0222011
; GENERAL INFORMATION:
; APPLICANT: Foote, Jefferson
; TITLE OF INVENTION: Super Humanized Antibodies
; FILE REFERENCE: 501231.02
; CURRENT APPLICATION NUMBER: PCT/US02/22011
; CURRENT FILING DATE: 2002-07-12
; PRIOR APPLICATION NUMBER: US 60/305,111
; PRIOR FILING DATE: 2001-07-12
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 121
; LENGTH: 119
; TYPE: PRT
; ORGANISM: Mus sp.
PCT-US02-22011-121

Query Match 100.0%; Score 23; DB 1; Length 119;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
|
|
|
|
Db 101 DYDA 104

RESULT 21

PCT-US02-22011-122
; Sequence 122, Application PC/TUS0222011
; GENERAL INFORMATION:
; APPLICANT: Foote, Jefferson
; TITLE OF INVENTION: Super Humanized Antibodies
; FILE REFERENCE: 501231.02
; CURRENT APPLICATION NUMBER: PCT/US02/22011
; CURRENT FILING DATE: 2002-07-12
; PRIOR APPLICATION NUMBER: US 60/305,111

; PRIOR FILING DATE: 2001-07-12
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 122
; LENGTH: 119
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: humanized anti-human glutamic acid decarboxylase (GAD65) Fab
PCT-US02-22011-122

Query Match 100.0%; Score 23; DB 1; Length 119;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
|
|
|
|
Db 101 DYDA 104

RESULT 22

US-10-194-975-121
; Sequence 121, Application US/10194975
; GENERAL INFORMATION:
; APPLICANT: Foote, Jefferson
; TITLE OF INVENTION: Super Humanized Antibodies
; FILE REFERENCE: 501231.01
; CURRENT APPLICATION NUMBER: US/10/194,975
; CURRENT FILING DATE: 2002-10-10
; PRIOR APPLICATION NUMBER: US 60/305,111
; PRIOR FILING DATE: 2001-07-12
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 121
; LENGTH: 119
; TYPE: PRT
; ORGANISM: Mus sp.
US-10-194-975-121

Query Match 100.0%; Score 23; DB 6; Length 119;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
|
|
|
|
Db 101 DYDA 104

RESULT 23

US-10-194-975-122
; Sequence 122, Application US/10194975
; GENERAL INFORMATION:
; APPLICANT: Foote, Jefferson
; TITLE OF INVENTION: Super Humanized Antibodies
; FILE REFERENCE: 501231.01
; CURRENT APPLICATION NUMBER: US/10/194,975
; CURRENT FILING DATE: 2002-10-10
; PRIOR APPLICATION NUMBER: US 60/305,111
; PRIOR FILING DATE: 2001-07-12
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 122
; LENGTH: 119
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: humanized anti-human glutamic acid decarboxylase (GAD65) Fab
US-10-194-975-122

Query Match 100.0%; Score 23; DB 6; Length 119;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
|||||
Db 101 DYDA 104

RESULT 24

US-10-264-237-1823
; Sequence 1823, Application US/10264237
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA131P1
; CURRENT APPLICATION NUMBER: US/10/264,237
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/16450
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: US 60/205,515
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 2876
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 1823
; LENGTH: 128
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-264-237-1823

Query Match 100.0%; Score 23; DB 6; Length 128;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
|||||
Db 39 DYDA 42

RESULT 25

PCT-US02-32727-5855
; Sequence 5855, Application PC/TUS0232727
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer
; APPLICANT: Skeiky, Yasir
; APPLICANT: Persing, David
; APPLICANT: Bhatia, Ajay
; APPLICANT: Maisonneuve, Jean Francois
; APPLICANT: Zhang, Yanni
; APPLICANT: Wang, Siqing
; APPLICANT: Jen, Shylian
; APPLICANT: Lodes, Michael
; APPLICANT: Benson, Darin
; APPLICANT: Jones, Robert
; APPLICANT: Carter, Darrick
; APPLICANT: Barth, Brenda
; APPLICANT: Douglass, John

US-10-264-237-1823
; Sequence 1823, Application US/10264237
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA131P1
; CURRENT APPLICATION NUMBER: US/10/264,237
; CURRENT FILING DATE: 2002-10-11
; NUMBER OF SEQ ID NOS: 30992
; SEQ ID NO 5855
; LENGTH: 130
; TYPE: PRT
; ORGANISM: Propionibacterium acnes

Query Match 100.0%; Score 23; DB 1; Length 130;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
|||||
Db 36 DYDA 39

US-10-057-498-5855
; Sequence 5855, Application US/10057498
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer
; APPLICANT: Skeiky, Yasir
; APPLICANT: Persing, David
; TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acne
; FILE REFERENCE: 210121.514
; CURRENT APPLICATION NUMBER: US/10/057,498
; CURRENT FILING DATE: 2001-04-20
; NUMBER OF SEQ ID NOS: 29212
; SEQ ID NO 5855
; LENGTH: 130
; TYPE: PRT
; ORGANISM: Propionibacterium acnes
US-10-057-498-5855

Query Match 100.0%; Score 23; DB 6; Length 130;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
|||||
Db 36 DYDA 39

RESULT 27

US-09-724-676-51020
; Sequence 51020, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 51020
; LENGTH: 145
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-51020

Query Match 100.0%; Score 23; DB 5; Length 145;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
|||||
Db 130 DYDA 133

RESULT 28

US-09-724-676A-51020
; Sequence 51020, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 51020
; LENGTH: 145
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676A-51020

Query Match 100.0%; Score 23; DB 5; Length 145;

```
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
    ||||
Db 130 DYDA 133

RESULT 29
US-09-950-084-4003
; Sequence 4003, Application US/09950084
; GENERAL INFORMATION:
; APPLICANT: George H. Shimer, Jr.
; APPLICANT: George H. Miller
; APPLICANT: Robert S. Hare
; APPLICANT: Karen J. Shaw
; TITLE OF INVENTION: Staphylococcus aureus Related Compositions and Methods
; FILE REFERENCE: 1034/1C363US2
; CURRENT APPLICATION NUMBER: US/09/950,084
; CURRENT FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: US 09/417,811
; PRIOR FILING DATE: 1999-10-14
; PRIOR APPLICATION NUMBER: US 09/353,718
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: US 09/266,557
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 09/266,556
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 09/266,555
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 09/266,542
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 09/266,541
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 09/037,934
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: US 09/036,720
; PRIOR FILING DATE: 1998-03-06
; PRIOR APPLICATION NUMBER: US 09/036,338
; PRIOR FILING DATE: 1998-03-06
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 7451
; SEQ ID NO 4003
; LENGTH: 159
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-950-084-4003

Query Match 100.0%; Score 23; DB 5; Length 159;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
    ||||
Db 117 DYDA 120

RESULT 30
US-10-264-237-1916
; Sequence 1916, Application US/10264237
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA131PI
; CURRENT APPLICATION NUMBER: US/10/264,237
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/16450
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: US 60/205,515
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 2876
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 1916
```

```
; LENGTH: 169
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-264-237-1916

Query Match 100.0%; Score 23; DB 6; Length 169;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
    ||||
Db 157 DYDA 160

RESULT 31
PCT-US02-32727-23164
; Sequence 23164, Application PC/TUS0232727
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer
; APPLICANT: Skeiky, Yasir
; APPLICANT: Persing, David
; APPLICANT: Bhatia, Ajay
; APPLICANT: Maisonneuve, Jean Francois
; APPLICANT: Zhang, Yanni
; APPLICANT: Wang, Siqing
; APPLICANT: Jen, Shyian
; APPLICANT: Lodes, Michael
; APPLICANT: Benson, Darin
; APPLICANT: Jones, Robert
; APPLICANT: Carter, Darrick
; APPLICANT: Barth, Brenda
; APPLICANT: Douglass, John
; TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of A
; FILE REFERENCE: 210121.514C1
; CURRENT APPLICATION NUMBER: PCT/US02/32727
; CURRENT FILING DATE: 2002-10-11
; NUMBER OF SEQ ID NOS: 30992
; SEQ ID NO 23164
; LENGTH: 170
; TYPE: PRT
; ORGANISM: Propioni acnes
PCT-US02-32727-23164

Query Match 100.0%; Score 23; DB 1; Length 170;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
    ||||
Db 159 DYDA 162

RESULT 32
US-10-057-498-23164
; Sequence 23164, Application US/10057498
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer
; APPLICANT: Skeiky, Yasir
; APPLICANT: Persing, David
; TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of A
; FILE REFERENCE: 210121.514
; CURRENT APPLICATION NUMBER: US/10/057,498
; CURRENT FILING DATE: 2001-04-20
; NUMBER OF SEQ ID NOS: 29212
; SEQ ID NO 23164
; LENGTH: 170
; TYPE: PRT
; ORGANISM: Propioni acnes
US-10-057-498-23164

Query Match 100.0%; Score 23; DB 6; Length 170;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```



```

; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-950-084-5911

Query Match      100.0%; Score 23; DB 5; Length 173;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
    ||||
Db 66 DYDA 69

RESULT 35
US-09-724-676-51015
; Sequence 51015, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 51015
; LENGTH: 175
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-51015

Query Match      100.0%; Score 23; DB 5; Length 175;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
    ||||
Db 160 DYDA 163

RESULT 36
US-09-724-676A-51015
; Sequence 51015, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 51015
; LENGTH: 175
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676A-51015

Query Match      100.0%; Score 23; DB 5; Length 175;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
    ||||
Db 160 DYDA 163

RESULT 37
US-09-950-084-3836
; Sequence 3836, Application US/09950084
; GENERAL INFORMATION:
; APPLICANT: George H. Shimer, Jr.
; APPLICANT: George H. Miller
; APPLICANT: Roberta S. Hare
; APPLICANT: Karen J. Shaw

```

```

QY 1 DYDA 4
    ||||
Db 159 DYDA 162

RESULT 33
US-10-276-774-2363
; Sequence 2363, Application US/10276774
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; APPLICANT: Tang, Y, Tom et al
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-030
; CURRENT APPLICATION NUMBER: US/10/276,774
; CURRENT FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 2700
; SOFTWARE: Custom
; SEQ ID NO 2363
; LENGTH: 171
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-276-774-2363

Query Match      100.0%; Score 23; DB 6; Length 171;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
    ||||
Db 146 DYDA 149

RESULT 34
US-09-950-084-5911
; Sequence 5911, Application US/09950084
; GENERAL INFORMATION:
; APPLICANT: George H. Shimer, Jr.
; APPLICANT: George H. Miller
; APPLICANT: Roberta S. Hare
; APPLICANT: Karen J. Shaw
; TITLE OF INVENTION: Staphylococcus aureus Related Compositions and Methods
; FILE REFERENCE: 1034/1C963US2
; CURRENT APPLICATION NUMBER: US/09/950,084
; CURRENT FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: US 09/417,811
; PRIOR FILING DATE: 1999-10-14
; PRIOR APPLICATION NUMBER: US 09/353,718
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: US 09/266,557
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 09/266,556
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 09/266,555
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 09/266,542
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 09/266,541
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 09/037,934
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: US 09/036,720
; PRIOR FILING DATE: 1998-03-06
; PRIOR APPLICATION NUMBER: US 09/036,338
; PRIOR FILING DATE: 1998-03-06
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 7451
; SEQ ID NO 5911
; LENGTH: 173

```

```
; TITLE OF INVENTION: Staphylococcus aureus Related Compositions and Methods
; FILE REFERENCE: 1034/IC963US2
; CURRENT APPLICATION NUMBER: US/09/950,084
; CURRENT FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: US 09/417,811
; PRIOR FILING DATE: 1999-10-14
; PRIOR APPLICATION NUMBER: US 09/353,718
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: US 09/266,557
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 09/266,556
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 09/266,555
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 09/266,542
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 09/266,541
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 09/037,934
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: US 09/036,720
; PRIOR FILING DATE: 1998-03-06
; PRIOR APPLICATION NUMBER: US 09/036,338
; PRIOR FILING DATE: 1998-03-06
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 7451
; SEQ ID NO 3836
; LENGTH: 175
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
; US-09-950-084-3836

Query Match 100.0%; Score 23; DB 5; Length 175;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
   ||||
Db 58 DYDA 61

RESULT 38
US-09-134-000C-4718
; Sequence 4718, Application US/09134000C
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4718
; LENGTH: 202
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
; US-09-134-000C-4718

Query Match 100.0%; Score 23; DB 5; Length 202;
Best Local Similarity 100.0%; Pred. No. 4.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
   ||||
Db 88 DYDA 91

RESULT 39
US-09-134-000C-4718
; Sequence 4718, Application US/09134000C
```

```
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4718
; LENGTH: 202
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
; US-09-134-000C-4718

Query Match 100.0%; Score 23; DB 5; Length 202;
Best Local Similarity 100.0%; Pred. No. 4.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
   ||||
Db 88 DYDA 91

RESULT 40
US-09-134-000C-3471
; Sequence 3471, Application US/09134000C
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3471
; LENGTH: 206
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
; US-09-134-000C-3471

Query Match 100.0%; Score 23; DB 5; Length 206;
Best Local Similarity 100.0%; Pred. No. 4.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
   ||||
Db 63 DYDA 66

RESULT 41
US-09-134-000C-3471
; Sequence 3471, Application US/09134000C
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3471
; LENGTH: 206
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
```

US-09-134-000C-3471

Query Match 100.0%; Score 23; DB 5; Length 206;
Best Local Similarity 100.0%; Pred. No. 4.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
||||
Db 63 DYDA 66

RESULT 42

US-09-724-676-67070
; Sequence 67070, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 67070
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-67070

Query Match 100.0%; Score 23; DB 5; Length 225;
Best Local Similarity 100.0%; Pred. No. 4.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
||||
Db 20 DYDA 23

RESULT 43

US-09-724-676-67071
; Sequence 67071, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 67071
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-67071

Query Match 100.0%; Score 23; DB 5; Length 225;
Best Local Similarity 100.0%; Pred. No. 4.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
||||
Db 20 DYDA 23

RESULT 44

US-09-724-676-67072
; Sequence 67072, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222

; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 67072
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-67072

Query Match 100.0%; Score 23; DB 5; Length 225;
Best Local Similarity 100.0%; Pred. No. 4.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
||||
Db 20 DYDA 23

RESULT 45

US-09-724-676-67073
; Sequence 67073, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 67073
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-67073

Query Match 100.0%; Score 23; DB 5; Length 225;
Best Local Similarity 100.0%; Pred. No. 4.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
||||
Db 20 DYDA 23

Search completed: February 6, 2003, 11:39:00
Job time : 15 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 6, 2003, 11:16:00 ; Search time 5.33333 Seconds
(without alignments)
31.107 Million cell updates/sec

Title: PAT943-1
Perfect score: 23
Sequence: 1 dyda 4

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues
Total number of hits satisfying chosen parameters: 319

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 100%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	23	100.0	85	1 YI31_MYCTU	Q50602 mycobacteri
2	23	100.0	91	1 YX57_MYCTU	O50386 mycobacteri
3	23	100.0	99	1 CU02_BACR	P80676 blaberus cr
4	23	100.0	103	1 SPI4_SOLTU	P58517 solanum tub
5	23	100.0	103	1 VG17_BPP22	P14109 bacterioph
6	23	100.0	105	1 NIFW_ANASP	O44149 anabaena sp
7	23	100.0	111	1 NIFW_SYN8	O07360 synecococc
8	23	100.0	114	1 SH3L_HUMAN	O75368 homo sapien
9	23	100.0	114	1 SH3L_MOUSE	Q91ju8 mus musculu
10	23	100.0	130	1 RS9_RALSO	Q8Y245 ralstonia s
11	23	100.0	131	1 COAT_BPPRR	P03616 bacterioph
12	23	100.0	141	1 RL25_YEAST	P04456 saccharomyc
13	23	100.0	143	1 HS18_STRAL	O53595 streptomyc
14	23	100.0	146	1 N7BM_CAEEL	Q9a2w7 caenorhabdi
15	23	100.0	146	1 RL2C_CAEEL	Q20647 caenorhabdi
16	23	100.0	147	1 RL2B_CAEEL	P48162 caenorhabdi
17	23	100.0	147	1 Y142_METJA	Q57607 methanococc
18	23	100.0	148	1 CB20_XENLA	P52299 xenopus lae
19	23	100.0	148	1 YF58_MYCTU	O10772 mycobacteri
20	23	100.0	154	1 RL2B_FRIAG	O22644 fritillaria
21	23	100.0	154	1 RL2B_TOBAC	Q07761 nicotiana t
22	23	100.0	156	1 CB20_HUMAN	P52298 homo sapien
23	23	100.0	156	1 GREU_UREPA	Q9pq17 ureaplasma
24	23	100.0	156	1 RISB_VIBCH	O9kpu4 vibrio chol
25	23	100.0	156	1 RL2B_HUMAN	P29316 homo sapien
26	23	100.0	158	1 SH3B_DROME	Q9nfp5 drosophila
27	23	100.0	160	1 FMA0_BACNO	P27688 bacteroides
28	23	100.0	160	1 FMA7_BACNO	P27691 bacteroides
29	23	100.0	160	1 GREU_MYCPN	P48019 mycoplasma
30	23	100.0	160	1 YEM0_YEAST	P40019 saccharomyc
31	23	100.0	161	1 FMA3_BACNO	P27689 bacteroides
32	23	100.0	161	1 GREU_MYCGE	P47524 mycoplasma
33	23	100.0	162	1 TPX_STRGC	P42366 streptococc

34	23	100.0	162	1 TPX_STRSA	P31308 streptococc
35	23	100.0	164	1 PTSN_HAEIN	P45072 haemophilus
36	23	100.0	170	1 GLB3_LUMTE	P11069 lumbricus t
37	23	100.0	171	1 Y175_STAAM	Q53719 staphylococ
38	23	100.0	177	1 OLFA_RANPI	P06910 rana pipien
39	23	100.0	186	1 SPI2_SOLTU	P58515 solanum tub
40	23	100.0	186	1 YZEE_STRCO	Q9k470 streptomyc
41	23	100.0	187	1 SPI1_SOLTU	P58514 solanum tub
42	23	100.0	197	1 WRBA_ECOLI	P30849 escherichia
43	23	100.0	197	1 WRBA_SALTI	Q8z7n9 salmonella
44	23	100.0	197	1 WRBA_SALTY	Q8zq40 salmonella
45	23	100.0	199	1 SOD1_HALNI	P09737 halobacteri

ALIGNMENTS

RESULT 1
YI31_MYCTU
ID YI31_MYCTU STANDARD; PRT: 85 AA.
AC Q50602:
01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein Rv1831.
GN Rv1831 OR MT1879.1 OR MTCYIA11.12C.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares R., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
complete genome sequence."
RL Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
laboratory strains."
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z78020; CAB01471.1; -;
DR EMBL; AE007046; -; NOT_ANNOTATED_CDS.
DR TIGR; MT1879.1; -;
DR TuberculList; Rv1831; -;
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 85 AA; 9696 MW; A20A383BE1D0D432 CRC64;

Query Match 100.0%; Score 23; DB 1; Length 85;

Best Local Similarity 100.0%; Pred. No. 58;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
|||||
Db 46 DYDA 49

RESULT 2
YX57_MYCTU
ID YX57_MYCTU STANDARD; PRT; 91 AA.
AC O50386;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein RV3357.
GN RV3357 OR MT3465 OR MTV004.14.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37Rv;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Church C., Harris D.,
Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
Davies K., Devlin K., Feltham T., Gentles S., Hamlin N., Holtroyd S.,
Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
complete genome sequence.";
RL Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
Kolony J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
Bishai W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
laboratory strains.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE PHD/YEPM FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; AL009198; CAA15742.1; -;
EMBL; AE007154; AAK47804.1; -;
TIGR; MT3465; -;
DR TubercuList; RV3357; -;
DR InterPro; IPR003756; DUF172.
DR Pfam; PF02604; DUF172; 1.
DR Hypothetical protein; Complete proteome.
SQ SEQUENCE 91 AA, 10194 MW, 544140972F329F0A CRC64;

Query Match 100.0%; Score 23; DB 1; Length 91;
Best Local Similarity 100.0%; Pred. No. 63;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
|||||
Db 42 DYDA 45

RESULT 3
CU02_BLACR
ID CU02_BLACR STANDARD; PRT; 99 AA.
AC P80676;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Cuticle protein 2 (BC-NCP2).
OS Blaberus craniifer.
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;
OC Blaberoidea; Blaberidae; Blaberus.
OX NCBI_TaxID=6982;
RN [1]
RP SEQUENCE.
RC TISSUE=Polyric caeca;
RX MEDLINE=97218697; PubMed=9066121;
RA Jensen U.G., Rothmann A., Skou L., Andersen S.O., Roepstorff P.,
Hoejrup P.;
RT "Cuticular proteins from the giant cockroach, Blaberus craniifer.";
RL Insect Biochem. Mol. Biol. 27:109-120(1997).
KW Structural protein; Cuticle; Repeat.
FT MOD_RES 1 1 PYROLIDONE CARBOXYLIC ACID.
FT DOMAIN 35 42 POLY-ALA.
FT REPEAT 16 31 1.
FT REPEAT 77 92 2.
SQ SEQUENCE 99 AA; 9867 MW; 690142670EA491BB CRC64;

Query Match 100.0%; Score 23; DB 1; Length 99;
Best Local Similarity 100.0%; Pred. No. 69;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
|||||
Db 44 DYDA 47

RESULT 4
SPI4_SOLTU
ID SPI4_SOLTU STANDARD; PRT; 103 AA.
AC P58517;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Serine protease inhibitor 4 (Pi-4) (Fragment).
OS Solanum tuberosum (Potato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4113;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Pentland squire; TISSUE=Tuber;
RX MEDLINE=92384955; PubMed=1515078;
RA Strukelj B., Pungercar J., Mesko P., Barlic-Maganja D., Gubensek F.,
Kregar I., Turk V.;
RT "Characterization of aspartic proteinase inhibitors from potato at
the gene cDNA and protein levels.";
RL Biol. Chem. Hoppe-Seyler 373:477-482(1992).
CC -!- FUNCTION: INHIBITOR OF SERINE PROTEASE. MAY PROTECT THE PLANT BY
INHIBITING PROTEASES OF INVADING ORGANISMS (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Vacuolar (By similarity).
CC -!- SIMILARITY: BELONGS TO THE LEGUMINOUS KUNITZ-TYPE INHIBITOR
FAMILY.
DR InterPro; IPR002160; Kunitz_legume.
DR ProDom; PD000891; Kunitz_legume; 1.
DR SMART; SM00452; STI; 1.
DR PROSITE; PS00283; SOYBEAN_KUNITZ; PARTIAL.
KW Serine protease inhibitor; Multigene family.
FT NON_TER 1 1
FT CHAIN <1 103 SERINE PROTEASE INHIBITOR 4.

```
FT ACT_SITE 25 26 REACTIVE BOND (FOR CHYMOTRYPSIN) (BY
FT DISULFID 56 73 SIMILARITY).
FT BY SIMILARITY.
SQ SEQUENCE 103 AA; 11349 MW; D328692B166B21C4 CRC64;

Query Match 100.0%; Score 23; DB 1; Length 103;
Best Local Similarity 100.0%; Pred. No. 72;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
Db 17 DYDA 20

RESULT 5
VG17_BPP22 STANDARD; PRT; 103 AA.
AC P14109;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Superinfection exclusion protein (Protein gp17).
GN 17.
OS Bacteriophage P22.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Podoviridae;
OC P22-like viruses.
OX NCBI_TaxID=10754;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89293845; PubMed=2738922;
RT "Samerjian A.V., Malloy D.C., Poteete A.R.;
RL J. Mol. Biol. 207:1-13(1989).
RN [2]
RP SEQUENCE FROM N.A.
RA Kropinski A.M.B., VanderByl C.S.;
RT "The completed sequence of genome of Salmonella phage P22.";
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: IN THE ABSENCE OF GP17, P22 IS SENSITIVE TO A
CC SUPERINFECTION EXCLUSION SYSTEM OF THE FEL-2 PROPAGE.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X15637; CAA33650.1; -.
DR EMBL; AF217253; AAF75018.1; -.
DR PIR; S04246; W1BP22.
SQ SEQUENCE 103 AA; 12184 MW; A8D12247123320E2C CRC64;

Query Match 100.0%; Score 23; DB 1; Length 103;
Best Local Similarity 100.0%; Pred. No. 72;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
Db 35 DYDA 38

RESULT 6
NIFW_ANASP STANDARD; PRT; 105 AA.
AC Q44149;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Nitrogenase stabilizing/protective protein nifw.
GN NIFW OR AL1433.
OS Anabaena sp. (strain PCC 7120).
```

```
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBI_TaxID=103690;
RN [1]
RP SEQUENCE FROM N.A.
RA Buikema W.J., Scappino L.A., Haselkorn R.;
RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21595285; PubMed=11759840;
RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
RA Watanabe A., Iriiguchi M., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA Nakazaki N., Shimpou S., Sugimoto M., Takazawa M., Yamada M.,
RA Yasuda M., Tabata S.;
RT "Complete genomic sequence of the filamentous nitrogen-fixing
RT cyanobacterium Anabaena sp. strain PCC 7120.";
RL DNA Res. 8:205-213(2001).
CC -1- FUNCTION: May protect the nitrogenase Fe-Mo protein from oxidative
CC damage (By similarity).
CC -1- SUBUNIT: Homotrimer; associates with nifD (By similarity).
CC -1- SIMILARITY: BELONGS TO THE NIFW FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U47055; AAA87952.1; -.
DR EMBL; AP003585; BAB73390.1; -.
DR InterPro; IPR004893; Nifw.
DR Pfam; PF03206; Nifw; 1
KW Nitrogen fixation; Complete proteome.
SQ SEQUENCE 105 AA; 12262 MW; 498DF9B8ADB8B95F CRC64;

Query Match 100.0%; Score 23; DB 1; Length 105;
Best Local Similarity 100.0%; Pred. No. 73;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
Db 24 DYDA 27

RESULT 7
NIFW_SYNPH STANDARD; PRT; 111 AA.
AC O07360;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Nitrogenase stabilizing/protective protein nifw.
GN NIFW.
OS Synechococcus sp. (strain PCC 8801 / RF-1) (Cyanothecae PCC 8801).
OC Bacteria; Cyanobacteria; Chroococcales; Cyanothecae.
OX NCBI_TaxID=41431;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99231861; PubMed=10217509;
RA Huang T.-C., Lin R.-F., Chu M.-K., Chen H.-M.;
RT "Organization and expression of nitrogen-fixation genes in the aerobic
RT nitrogen-fixing unicellular cyanobacterium Synechococcus sp. strain
RT RF-1.";
RL Microbiology 145:743-753(1999).
CC -1- FUNCTION: May protect the nitrogenase Fe-Mo protein from oxidative
CC damage (By similarity).
CC -1- SUBUNIT: Homotrimer; associates with nifD (By similarity).
CC -1- SIMILARITY: BELONGS TO THE NIFW FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
```

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement. (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL; AF003700; AAC35197.1; -;
 DR InterPro; IPR004893; Nifw.
 DR Pfam; PF03206; Nifw; 1.
 KW Nitrogen fixation.
 SQ SEQUENCE 111 AA; 12931 MW; C85B39697EFC310E CRG64;

Query Match 100.0%; Score 23; DB 1; Length 111;
 Best Local Similarity 100.0%; Pred. No. 78;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
 DB 27 DYDA 30

RESULT 8

SH3L_HUMAN
 ID SH3L_HUMAN STANDARD; PRT; 114 AA.
 AC 075368;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE SH3 domain-binding glutamic acid-rich-like protein.
 GN SH3BGR.
 OS Homo sapiens (Human).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC Tissue=Brain;
 RX MEDLINE=98308110; PubMed=9642120;
 RA Egeo A., Mazzocco M., Arrigo P., Vidal-Taboada J.M., Oliva R.,
 RA Pirola B., Giglio S., Rasore-Quartino A., Scartezzini P.;
 RT "Identification and characterization of a new human gene encoding a
 RT small protein with high homology to the proline-rich region of the
 RT SH3BGR gene.";
 RL Biochem. Biophys. Res. Commun. 247:302-306(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC Tissue=Bone marrow;
 RA Strausberg R.;
 RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
 CC -!- TISSUE SPECIFICITY: UBIQUITOUS.
 CC -!- SIMILARITY: BELONGS TO THE SH3BGR FAMILY.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement. (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL; AF042081; AAC27445.1; -;
 DR EMBL; BC016709; AAH16709.1; -;
 DR Genew; HGNC:10823; SH3BGR.
 DR MIM; 300190; -;
 KW SH3-binding.
 FT SITE 61 67 SH3-BINDING (POTENTIAL).
 SQ SEQUENCE 114 AA; 12774 MW; D72AC2A095F1AA8B CRG64;

Query Match 100.0%; Score 23; DB 1; Length 114;
 Best Local Similarity 100.0%; Pred. No. 80;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4

DB 78 DYDA 81
 IIII

RESULT 9

SH3L_MOUSE
 ID SH3L_MOUSE STANDARD; PRT; 114 AA.
 AC 09JJU8;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE SH3 domain-binding glutamic acid-rich-like protein.
 GN SH3BGR.
 OS Mus musculus (Mouse).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC Tissue=Liver;
 RA Scartezzini P., Egeo A., Mazzocco M.;
 RT "Cloning the mouse homolog of the human SH3BGR gene.";
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: BELONGS TO THE SH3BGR FAMILY.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement. (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL; AJ272504; CAB76919.1; -;
 DR MGD; MGI:1930849; Sh3bgrl.
 KW SH3-binding.
 FT SITE 61 67 SH3-BINDING (POTENTIAL).
 SQ SEQUENCE 114 AA; 12811 MW; D7298AF8F5BFA43F CRG64;

Query Match 100.0%; Score 23; DB 1; Length 114;
 Best Local Similarity 100.0%; Pred. No. 80;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
 DB 78 DYDA 81
 IIII

RESULT 10

RS9_RALSO
 ID RS9_RALSO STANDARD; PRT; 130 AA.
 AC 08Y245;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DE 30S ribosomal protein S9.
 GN RPS1 OR RSC0491 OR RS05034.
 OS Ralstonia solanacearum (Pseudomonas solanacearum).
 CC Bacteria; Proteobacteria; beta subdivision; Ralstonia group;
 CC Ralstonia.
 OX NCBI_TaxID=305;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=GMI1000;
 RX MEDLINE=21681879; PubMed=11823852;
 RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
 RA Ariat M., Billault A., Brottier P., Camus J.C., Catolico L.,
 RA Chandler M., Choise N., Claudel-Renard C., Cunnac S., Demange N.,
 RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schliex T.,
 RA Siquier P., Thebaud P., Whalen M., Wincker P., Levy M.,
 RA Weissenbach J., Boucher C.A.;
 RT "Genome sequence of the plant pathogen Ralstonia solanacearum.";
 RL Nature 415:497-502(2002).

CC -!- SIMILARITY: BELONGS TO THE S9P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AL646059; CAD14019.1; -;
DR InterPro: IPR000754; Ribosomal_S9.
DR Pfam: PF00380; Ribosomal_S9; 1.
DR PROSITE: PS00360; RIBOSOMAL_S9; 1.
KW Ribosomal protein; Complete proteome.
SQ SEQUENCE 130 AA; 14336 MW; A539A49A953DE1A5 CRC64;

Query Match 100.0%; Score 23; DB 1; Length 130;
Best Local Similarity 100.0%; Pred. No. 92;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DYDA 4
Db 89 DYDA 92

RESULT 11
COAT_BPBR STANDARD; PRT; 131 AA.
AC P03616;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last annotation update)
DE Coat protein. PRRL.
OS Bacteriophage PRRL.
CC Viruses; ssRNA positive-strand viruses, no DNA stage; Leviviridae;
CC unclassified Leviviridae.
OX NCBI_TaxID=12024;
RN [1]
RP SEQUENCE.
RX MEDLINE=79148387; PubMed=107028;
RA Dhaese P., Vandekerckhove J., van Montagu M.;
RT "The primary structure of the coat protein of the broad-host-range
RT RNA bacteriophage PRRL."
RL Eur. J. Biochem. 94:375-386(1979).
CC -!- FUNCTION: FORMS THE PHAGE SHELL; BINDS TO THE PHAGE RNA.
DR PIR: A04225; VCBPp1.
DR InterPro: IPR002703; Levi_coat.
DR Pfam: PF01819; Levi_coat; 1.
KW Coat protein; RNA-binding.
SQ SEQUENCE 131 AA; 14535 MW; E7E639E1E50FC612 CRC64;

Query Match 100.0%; Score 23; DB 1; Length 131;
Best Local Similarity 100.0%; Pred. No. 93;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DYDA 4
Db 92 DYDA 95

RESULT 12
RL25_YEAST STANDARD; PRT; 141 AA.
ID RL25_YEAST
AC P04456;
DT 13-AUG-1987 (Rel. 05, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE 60S ribosomal protein L25 (YL25) (RP61L).
GN RPL25 OR YOL127W.
OS Saccharomyces cerevisiae (Baker's yeast).
CC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
CC Saccharomycetales; Saccharomycetaceae; Saccharomyces.

OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Carlsbergensis;
RX MEDLINE=8504125; PubMed=6091033;
RA Leer R.J., van Raamsdonk-Duin M.M.C., Hagendoorn M.J.M., Mager W.H.,
RA Planta R.J.;
RT "Structural comparison of yeast ribosomal protein genes."
RL Nucleic Acids Res. 12:6685-6700(1984).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c / FY1679;
RX MEDLINE=97051588; PubMed=8896265;
RA Casamayor A., Khalid H., Balcells L., Aldea M., Casas C.,
RA Herrero E., Arino J.;
RT "Sequence analysis of a 13.4 kbp fragment from the left arm of
RT chromosome XV reveals a malate dehydrogenase gene, a putative Ser/Thr
RT protein kinase, the ribosomal L25 gene and four new open reading
RT frames."
RL Yeast 12:1013-1020(1996).
RN [3]
RP SEQUENCE OF 1-20.
RX MEDLINE=92184799; PubMed=1544921;
RA Takakura H., Tsunasawa S., Miyagi M., Warner J.R.;
RT "NH2-terminal acetylation of ribosomal proteins of Saccharomyces
RT cerevisiae."
RL J. Biol. Chem. 267:5442-5445(1992).
CC -!- FUNCTION: THIS PROTEIN BINDS TO A SPECIFIC REGION ON THE 26S RRNA.
CC -!- SIMILARITY: BELONGS TO THE L23P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U41293; AAC49465.1; -;
DR EMBL: X01014; CAA25506.1; -;
DR PIR: A02820; R5BY25.
DR SGD: S0005487; RPL25.
DR InterPro: IPR001014; Ribosomal_L23.
DR Pfam: PF00276; Ribosomal_L23; 1.
DR PROSITE: PD001141; Ribosomal_L23; 1.
DR RIBOSOMAL_L23; 1.
KW Ribosomal protein; rRNA-binding.
FT INIT_MET 0
FT CONFLICT 105 111 DVLKVT -> NI (IN REF. 1).
SQ SEQUENCE 141 AA; 15626 MW; C53ED3DA4A9D219A CRC64;

Query Match 100.0%; Score 23; DB 1; Length 141;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DYDA 4
Db 128 DYDA 131

RESULT 13
HS18_STRAL STANDARD; PRT; 143 AA.
ID HS18_STRAL
AC Q53595;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE 18 kDa heat shock protein (HSP 18).
GN HSP18.
OS Streptomyces albus G.
CC Bacteria; Actinobacteria; Actinobacteriales; Streptomycetaceae; Streptomyces.
CC Actinomycetales; Streptomycetaceae; Streptomyces.

```
OX NCBI_TaxID=1962;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=95286476; PubMed=7768794;
RA Servant P., Mazodier P.;
RT "Characterization of Streptomyces albus 18-kilodalton heat shock-
RL responsive protein.";
RL J. Bacteriol. 177:2998-3003(1995).
CC -|- INDUCTION: BY HEAT SHOCK.
CC -|- SIMILARITY: BELONGS TO THE SMALL HEAT SHOCK PROTEIN (HSP20)
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U17419; AAA86472.1; -.
DR InterPro: IPR002068; Hsp20.
DR Pfam: PF00011; HSP20; 1.
DR PROSITE: PS01031; HSP20; 1.
KW Heat shock.
SQ SEQUENCE 143 AA; 16107 MW; DCB8FA0DE154B197 CRC64;

Query Match 100.0%; Score 23; DB 1; Length 143;
Best Local Similarity 100.0%; Pred. NO. 1.le+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
Db 109 DYDA 112
|||||

RESULT 14
N7BM_CAEEL
ID N7BM_CAEEL STANDARD; PRT; 146 AA.
AC Q9NZM7.
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Putative NADH-ubiquinone oxidoreductase subunit B17.2 (EC 1.6.5.3)
DE (EC 1.6.99.3) (Complex I-B17.2) (CI-B17.2).
GN Y94H6A.8.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Bristol N2;
RA Waterston R.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
-|- FUNCTION: TRANSFER OF ELECTRONS FROM NADH TO THE RESPIRATORY
CC CHAIN. THE IMMEDIATE ELECTRON ACCEPTOR FOR THE ENZYME IS
CC BELIEVED TO BE UBIQUINONE (BY SIMILARITY).
CC -|- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
CC -|- CATALYTIC ACTIVITY: NADH + acceptor = NAD(+) + reduced acceptor.
CC -|- SUBUNIT: COMPLEX I IS COMPOSED OF ABOUT 41 DIFFERENT SUBUNITS (BY
CC SIMILARITY).
CC -|- SUBCELLULAR LOCATION: Mitochondrial inner membrane; matrix side
CC (By similarity).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
```

```
DR EMBL: AC024876; AAF60889.1; ALT_INIT.
DR WormPep; Y94H6A.8; CE23630.
KW Hypothetical protein; Oxidoreductase; NAD; Ubiquinone; Mitochondrion.
SQ SEQUENCE 146 AA; 17073 MW; 0A8549D622977786 CRC64;

Query Match 100.0%; Score 23; DB 1; Length 146;
Best Local Similarity 100.0%; Pred. NO. 1.le+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
Db 76 DYDA 79
|||||

RESULT 15
RL2C_CAEEL
ID RL2C_CAEEL STANDARD; PRT; 146 AA.
AC Q20647.
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Probable 60S ribosomal protein L23A 2.
GN F52B5.6.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Bristol N2;
RA Kershaw J.;
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
-|- FUNCTION: THIS PROTEIN BINDS TO A SPECIFIC REGION ON THE 26S RRNA
CC (BY SIMILARITY).
CC -|- SIMILARITY: BELONGS TO THE L23P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Z75541; CAA99858.1; -.
DR WormPep; F52B5.6; CE05721.
DR InterPro: IPR001014; Ribosomal_L23.
DR Pfam: PF00276; Ribosomal_L23; 1.
DR PRODOM: PD001141; Ribosomal_L23; 1.
DR PROSITE: PS00050; RIBOSOMAL_L23; 1.
KW Ribosomal protein; rRNA-binding.
SQ SEQUENCE 146 AA; 16283 MW; 3380545260EA88C7 CRC64;

Query Match 100.0%; Score 23; DB 1; Length 146;
Best Local Similarity 100.0%; Pred. NO. 1.le+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
Db 133 DYDA 136
|||||

RESULT 16
RL2B_CAEEL
ID RL2B_CAEEL STANDARD; PRT; 147 AA.
AC P48162.
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE Probable 60S ribosomal protein L23A 1.
GN F55D10.2.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
```

```
CC Rhabdittidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Leinbach D.;
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: THIS PROTEIN BINDS TO A SPECIFIC REGION ON THE 26S RRNA
CC (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE L23P FAMILY OF RIBOSOMAL PROTEINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U40948; AAA81728.1; -.
DR WormPep; F55D10.2; CF02777.
DR InterPro; IPR001014; Ribosomal_L23.
DR Pfam; PF00276; Ribosomal_L23; 1.
DR ProDom; PD001141; Ribosomal_L23; 1.
DR PROSITE; PS00050; RIBOSOMAL_L23; 1.
KW Ribosomal protein; rRNA-binding.
SQ SEQUENCE 147 AA; 16696 MW; 1F3FF0088C453955 CRC64;

Query Match 100.0%; Score 23; DB 1; Length 147;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
DB 134 DYDA 137

RESULT 17
Y142_METJA
ID Y142_METJA STANDARD; PRT; 147 AA.
AC Q57607;
DT 01-NOV-1997 (Rel. 35, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein MJ0142.
GN MJ0142
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcaceae; Methanocaldococcus.
OX NCBI_TaxID=2130;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., FitzGerald L.M., Clayton R.A., Gocayne J.D.,
RA Sutton G.G., Blake J.A., Fritch J.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA Scott J.L., Geoghegan N.S.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii."
RL Science 273:1058-1073(1996).
CC -!- SIMILARITY: STRONG, TO M.JANNASCHII M0605.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC
```

```
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U67471; AAB98125.1; -.
DR TIGR; MJ0142; -.
DR InterPro; IPR002806; DUF103.
DR Pfam; PF01953; DUF103; 1.
DR ProDom; PD015886; DUF103; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 147 AA; 17592 MW; 707D67BF058E065C CRC64;

Query Match 100.0%; Score 23; DB 1; Length 147;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
DB 102 DYDA 105

RESULT 18
CB20_XENLA
ID CB20_XENLA STANDARD; PRT; 148 AA.
AC PS2299;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 20 kDa nuclear cap binding protein (NCBP) (CBP20) (Fragment).
GN CBP20.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney epithelium;
RX MEDLINE=95379956; PubMed=7651522;
RA Izaurralde E.L., Lewis J., Gamberi C., Jarmolowski A.,
RA McGuigan C., Mattaj J.W.;
RT "A cap-binding protein complex mediating U snRNA export.";
RL Nature 376:709-712(1995).
CC -!- FUNCTION: INVOLVED IN MEDIATING U SNRNA EXPORT FROM THE NUCLEUS.
CC BINDS TO 5'CAPPED MRNA.
CC -!- SUBUNIT: THE CAP-BINDING PROTEIN (CBC) COMPLEX IS AN HETERODIMER
CC OF CBP80 AND CBP20.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X84788; CAA59259.1; -.
DR HSP; P09651; 1XAL.
DR InterPro; IPR000504; RNA_rec_mot.
DR Pfam; PF00076; rrm; 1.
DR SMART; SM00360; RRM; 1.
DR PROSITE; PS0102; RRM; 1.
DR PROSITE; PS00030; RRM_RNP_1; FALSE_NEG.
KW Nuclear protein; RNA-binding.
FT NON_TER 1
FT DOMAIN 32 110 RNA-BINDING (RRM).
SQ SEQUENCE 148 AA; 16920 MW; 007CDACABDC89A8C CRC64;

Query Match 100.0%; Score 23; DB 1; Length 148;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
```

```
Db 133 DYDA 136
||||
RESULT 19
YF58_MYCTU
ID YF58_MYCTU STANDARD; PRT; 148 AA.
AC Q10772;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein Rv1558.
GN Rv1558 OR MT1609 OR MTCY48.07C.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares R., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RA "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RL Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: TO M.TUBERCULOSIS.RV1261C.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
DR EMBL; Z74020; CAA98333.1; -
DR EMBL; AE007027; AAK45876.1; -
DR TIGR; MT1609; -
DR TubercuList; Rv1558; -
DR InterPro; IPR004378; Mtu_fam_11.
DR TIGRFAMs; TIGR00026; Mtu_fam_11; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 148 AA; 16347 MW; A9F98E90398719B8 CRC64;

Query Match 100.0%; Score 23; DB 1; Length 148;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
Db 100 DYDA 103
||||
RESULT 20
RL2B_FRIAG
ID RL2B_FRIAG STANDARD; PRT; 154 AA.
AC Q22644;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE 60S ribosomal protein L23A.
GN RPL23A.
OS Fritillaria agrestis.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Liliales; Liliaceae;
OC Fritillaria.
OX NCBI_TaxID=64177;
RN [1]
RP SEQUENCE FROM N.A.
RA Prabhavalkar D.S., Baysdorfer C.;
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: THIS PROTEIN BINDS TO A SPECIFIC REGION ON THE 26S RRNA
CC (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE L23P FAMILY OF RIBOSOMAL PROTEINS.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
DR EMBL; AF031542; AAB6852.1; -
DR InterPro; IPR001014; Ribosomal_L23.
DR Pfam; PF00276; Ribosomal_L23; 1.
DR ProDom; PD001141; Ribosomal_L23; 1.
DR PROSITE; PS00050; RIBOSOMAL_L23; 1.
KW Ribosomal protein; rRNA-binding.
SQ SEQUENCE 154 AA; 17350 MW; 52522CE2D969A52F CRC64;

Query Match 100.0%; Score 23; DB 1; Length 154;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
Db 141 DYDA 144
||||
RESULT 21
RL2B_TOBAC
ID RL2B_TOBAC STANDARD; PRT; 154 AA.
AC Q07761;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE 60S ribosomal protein L23A (L25).
GN RPL23A OR RPL25.
OS Nicotiana tabacum (Common tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=4097;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=94355650; PubMed=8075394;
RA Gao J., Kim S.R., Chung Y.Y., Lee J.M., An G.;
RT "Developmental and environmental regulation of two ribosomal protein
RT genes in tobacco.";
RL Plant Mol. Biol. 25:761-770(1994).
CC -!- FUNCTION: THIS PROTEIN BINDS TO A SPECIFIC REGION ON THE 26S RRNA
CC (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE L23P FAMILY OF RIBOSOMAL PROTEINS.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; L18908; AA53296.1; -
DR InterPro; IPR001014; Ribosomal_L23.
DR Pfam; PF00276; Ribosomal_L23; 1.
DR ProDom; PD001141; Ribosomal_L23; 1.
DR PROSITE; PS00050; Ribosomal_L23; 1.
KW Ribosomal protein; rRNA-binding.
SQ SEQUENCE 154 AA; 17281 MW; 84A845DF1161067E CRC64;

Query Match 100.0%; Score 23; DB 1; Length 154;
Best Local Similarity 100.0%; Pred. No. 1.le+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DYDA 4
|
|
|
|
DB 141 DYDA 144

RESULT 22
CB20_HUMAN
ID CB20_HUMAN STANDARD; PRT; 156 AA.
AC P52298; Q14924;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE 20 kDa nuclear cap binding protein (NCBP 20 kDa subunit) (CBP20) (NCBP
DE interacting protein 1) (NIP1).
GN NCBP2 OR CBP20.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 9-25 AND 113-145.
RX MEDLINE=95379956; PubMed=7651522;
RA Izaurralde E., Lewis J., Gambieri C., Jarmolowski A.,
RA McGuigan C., Mattaj A.W.;
RT "A cap-binding protein complex mediating U snRNA export.";
RL Nature 376:709-712(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Cervical carcinoma;
RX MEDLINE=96017765; PubMed=7478990;
RA Kataoka N., Ohno M., Mada I., Shimura Y.;
RT "Identification of the factors that interact with NCBP, an 80 kDa
RT nuclear cap binding protein.";
RL Nucleic Acids Res. 23:3638-3641(1995).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Cervix;
RA Strausberg R.;
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: INVOLVED IN MEDIATING U SNRNA EXPORT FROM THE NUCLEUS.
CC BINDS TO 5'CAPPED MRNA.
CC -1- SUBUNIT: THE CAP-BINDING PROTEIN (CBC) COMPLEX IS AN HETERODIMER
CC OF CBP80 AND CBP20.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; X84157; CAA58962.1; -
DR EMBL; D59253; BAA09599.1; -

DR EMBL; BC001255; AAH01255.1; -
DR HSSP; P11940; 1CVJ.
DR Genew; HGNC:7659; NCBP2.
DR MIM; 605133; -
DR InterPro; IPR000504; RNA_rec_mot.
DR Pfam; PF00076; rrm; 1.
DR SMART; SM00360; RRM; 1.
DR PROSITE; PS0102; RRM; 1.
DR PROSITE; PS00030; RRM_RNP_1; FALSE_NEG.
KW Nuclear protein; RNA-binding.
FT DOMAIN 40 118 RNA-BINDING (RRM).
FT CONFLICT 97 97 A -> S (IN REF. 2).
SQ SEQUENCE 156 AA; 18001 MW; B6C94F3182A2CC3D CRC64;

Query Match 100.0%; Score 23; DB 1; Length 156;
Best Local Similarity 100.0%; Pred. No. 1.le+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DYDA 4
|
|
|
|
DB 141 DYDA 144

RESULT 23
GREA_UREPA
ID GREA_UREPA STANDARD; PRT; 156 AA.
AC Q9PQ17; -
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Transcription elongation factor greA (Transcript cleavage factor
DE greA).
DE GREA.
GN GREA OR UU304.
OS Ureaplasma parvum (Ureaplasma urealyticum biotype 1).
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Ureaplasma.
OX NCBI_TaxID=134821;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Serovar 3;
RX MEDLINE=20500219; PubMed=11048724;
RA Glass J.I., Lefkowitz E.J., Glass J.S., Heiner C.R., Chen E.Y.,
RA Cassell G.H.;
RT "The complete sequence of the mucosal pathogen Ureaplasma
RT urealyticum.";
RL Nature 407:757-762(2000).
CC -1- FUNCTION: NECESSARY FOR EFFICIENT RNA POLYMERASE TRANSCRIPTION
CC ELONGATION PAST TEMPLATE-ENCODED ARRESTING SITES. THE ARRESTING
CC SITES IN DNA HAVE THE PROPERTY OF TRAPPING A CERTAIN FRACTION
CC OF ELONGATING RNA POLYMERASES THAT PASS THROUGH, RESULTING IN
CC LOCKED TERNARY COMPLEXES. CLEAVAGE OF THE NASCENT TRANSCRIPT BY
CC CLEAVAGE FACTORS SUCH AS GREA OR GREB ALLOWS THE RESUMPTION OF
CC ELONGATION FROM THE NEW 3' TERMINUS. GREA RELEASES SEQUENCES OF
CC 2 TO 3 NUCLEOTIDES (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE GREA/GREB FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; AE002128; AAF30713.1; -
DR HSSP; P21346; 1GRJ.
DR InterPro; IPR001437; Grea_Greb.
DR Pfam; PF01272; Grea_Greb; 1.
DR Pfam; PF03449; Grea_Greb_N; 1.
DR ProDom; PD004918; Grea_Greb; 1.
DR PROSITE; PS00829; GREAB_1; 1.
DR PROSITE; PS00830; GREAB_2; FALSE_NEG.
KW Transcription regulation; DNA-binding; Coiled coil; Complete proteome.
FT DOMAIN 7 24 COILED COIL (POTENTIAL).

```
FT DOMAIN 42 84 COILED COIL (POTENTIAL).
SQ SEQUENCE 156 AA; 17767 MW; 86BEA8EC07C96461 CRC64;

Query Match 100.0%; Score 23; DB 1; Length 156;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
Db 47 DYDA 50

RESULT 24
RISB_VIBCH STANDARD; PRT; 156 AA.
ID RISB_VIBCH
AC Q9KPU4;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE 6,7-dimethyl-8-ribityllumazine synthase (EC 2.5.1.9) (DMRL synthase)
DE (lumazine synthase) (Riboflavin synthase beta chain).
GN RIBH OR VC2268.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OC NCBI_TaxID=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=El Tor N16961 / Serotype O1;
RX MEDLINE=20406833; PubMed=10952301;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT cholerae."
RL Nature 406:477-483(2000).
CC
CC -!- FUNCTION: Riboflavin synthase, is a bifunctional enzyme complex
CC catalyzing the formation of riboflavin from 5-amino-6-(1'-D)-
CC ribityl-amino-2,4(1H,3H)-pyrimidinone and L-3,4-dihydroxy-2-
CC butanone-4-phosphate via 6,7-dimethyl-8-lumazine. The beta subunit
CC catalyzes the condensation of 5-amino-6-(1'-D)-ribityl-amino-
CC 2,4(1H,3H)-pyrimidinone with L-3,4-dihydroxy-2-butanone-4-
CC phosphate yielding 6,7-dimethyl-8-lumazine (by similarity).
CC -!- CATALYTIC ACTIVITY: 2 6,7-dimethyl-8-(1-D-ribityl)lumazine =
CC riboflavin + 4-(1-D-ribitylamino)-5-amino-2,6-dihydroxypyrimidine.
CC -!- PATHWAY: Riboflavin biosynthesis; last step.
CC -!- SIMILARITY: BELONGS TO THE DMRL SYNTHASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF004298; AAF95412.1; ALT_INIT.
DR HSSP; P11998; IRVV.
DR TIGR; VC2268; -.
DR InterPro; IPR002180; DMRL_synthase.
DR Pfam; PF00885; DMRL_synthase; 1.
DR ProDom; PD003664; DMRL_synthase; 1.
DR TIGRFAMs; TIGR00114; ribh; 1.
KW Riboflavin biosynthesis; transferase; Complete proteome.
SQ SEQUENCE 156 AA; 16433 MW; E7E9BAC14386DBE4 CRC64;

Query Match 100.0%; Score 23; DB 1; Length 156;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
Db 72 DYDA 75

RESULT 25
RL2B_HUMAN STANDARD; PRT; 156 AA.
ID RL2B_HUMAN
AC P29316; P39024; Q92774;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE 60S ribosomal protein L23a.
GN RPL23A.
OS Homo sapiens (Human),
OS Mus musculus (Mouse), and
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OC NCBI_TaxID=9606, 10090, 10116;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=Human; TISSUE=Melanoma;
RA Jiang H., Lin J., Tao J., Fisher P.B.;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=Human; TISSUE=Liver;
RA Reddy K.B., Lin C.W., Howe P.H.;
RL Submitted (SEP-1993) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC SPECIES=Human;
RX MEDLINE=96269983; PubMed=8662070;
RA Fan W., Cai W., Parimoo S., Lennon G.G., Weissman S.M.;
RT "Identification of seven new human MHC class I region genes around
RT the HLA-F locus."
RL Immunogenetics 44:97-103(1996).
RN [4]
RP SEQUENCE FROM N.A.
RC SPECIES=Human;
RX MEDLINE=98086480; PubMed=9417910;
RA Fan W., Christensen M., Eichler E., Zhang X., Lennon G.;
RT "Cloning, sequencing, gene organization, and localization of the
RT human ribosomal protein RPL23A gene."
RL Genomics 46:234-239(1997).
RN [5]
RP SEQUENCE FROM N.A.
RC SPECIES=Human; TISSUE=Skin;
RA Strausberg R.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE OF 10-156 FROM N.A.
RC SPECIES=Human; TISSUE=Liver;
RA Ronaldo M., Soares M.B.;
RL Submitted (XXX-1993) to the EMBL/GenBank/DBJ databases.
RN [7]
RP SEQUENCE FROM N.A.
RC SPECIES=Mouse; TISSUE=Liver;
RA Strausberg R.;
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
RN [8]
RP SEQUENCE FROM N.A.
RC SPECIES=Rat; STRAIN=Sprague-Dawley; TISSUE=Liver;
RX MEDLINE=93155091; PubMed=8428950;
RA Suzuki K., Wool I.G.;
RT "The primary structure of rat ribosomal protein L23a. The application
RT of homology search to the identification of genes for mammalian and
RT yeast ribosomal proteins and a correlation of rat and yeast ribosomal
RT proteins."
RL J. Biol. Chem. 268:2755-2761(1993).
CC -!- FUNCTION: THIS PROTEIN BINDS TO A SPECIFIC REGION ON THE 26S RRNA
CC (BY SIMILARITY).
```

```

CC -!- SIMILARITY: BELONGS TO THE L23P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U43701; AA03210.1; -
DR EMBL: U02032; AA03341.1; -
DR EMBL: U37230; AA017510.1; -
DR EMBL: AF001689; AAC51934.1; -
DR EMBL: BC014459; AAH14459.1; -
DR EMBL: L13799; AAA35681.1; -
DR EMBL: BC029892; AAH29892.1; -
DR EMBL: X65228; CAA46336.1; -
DR PIR: S19992; R3RT3A.
DR SWISS-2DPAGE: P29316; HUMAN.
DR GENE: HGNC:10317; RPL23A.
DR MIM: 602326; -
DR InterPro: IPR001014; Ribosomal_L23.
DR Pfam: PF00276; Ribosomal_L23; 1.
DR ProDom: PD001141; Ribosomal_L23; 1.
DR PROSITE: PS00050; RIBOSOMAL_L23; 1.
DR Ribosomal protein; rRNA-binding.
KW CONFLICT 1 2 MA -> IP (IN REF. 2).
FT CONFLICT 78 78 K -> N (IN REF. 3).
FT CONFLICT 110 110 K -> S (IN REF. 3).
SQ SEQUENCE 156 AA; 17695 MW; 3980E77B47FAB70E CRC64;

Query Match 100.0%; Score 23; DB 1; Length 156;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DYDA 4
Db 143 DYDA 146

RESULT 26
SH3B_DROME STANDARD; PRT; 158 AA.
AC Q9NFP5; Q9V5S8;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE SH3 domain-binding glutamic acid-rich protein homolog (SH3BGR
DE protein).
DE SH3BGR OR CG8582.
GN Drosophila melanogaster (Fruit fly).
OS Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC Tissue=Embryo;
RA Scarszini P., Egeo A., Mazzocco M.;
RT "Cloning of the Drosophila Sh3bgr gene homolog of human Sh3bgr gene.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkeley;
RX MEDLINE=20196006; PubMed=107311132;
RA Adams M.D., Cealniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,

```

```

RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
RA Balles R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Bottier P.,
RA Burtis J.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA De Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fosler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai X.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Zheng L.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
CC -!- SIMILARITY: BELONGS TO THE SH3BGR FAMILY.
CC -!- CAUTION: Ref.2 sequence differs from that shown due to erroneous
CC gene model prediction.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AJ272505; CAB76915.1; -
DR EMBL: AE003559; AAF50570.1; ALT_SEQ.
DR FlyBase: FBgn0035772; CG8582.
KW SH3-binding.
FT SITE 67 73 SH3-BINDING (POTENTIAL).
FT CONFLICT 13 14 NK -> LL (IN REF. 1).
FT CONFLICT 20 21 QQ -> HE (IN REF. 1).
FT CONFLICT 64 64 D -> V (IN REF. 1).
SQ SEQUENCE 158 AA; 17488 MW; 9DEF79C3D48CA463 CRC64;

Query Match 100.0%; Score 23; DB 1; Length 158;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DYDA 4
Db 84 DYDA 87

RESULT 27
FMA0_BACNO STANDARD; PRT; 160 AA.
ID FMA0_BACNO
AC P27688;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 01-AUG-1992 (Rel. 23, Last annotation update)
DE Fimbrial protein precursor (Pilin) (Serogroup B1/AC20).
GN FIMA.
OS Bacteroides nodosus (Dichelobacter nodosus).

```

```

OC Bacteria; Proteobacteria; gamma subdivision; Cardiobacteriaceae;
OX Dichelobacter.
OX NCBI_TaxID=870;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Serogroup B1 isolate AC20;
RX MEDLINE=91216435; PubMed=1673665;
RA Billington S.J., Rood J.I.;
RT "sequence of fibrial subunit-encoding genes from virulent and benign
RL isolates of Dichelobacter (Bacteroides) nodosus.";
RL Gene 99:115-119(1991).
CC -1- SUBUNIT: THE PILI ARE POLAR FLEXIBLE FILAMENTS OF ABOUT 5.4
CC NANOMETERS DIAMETER AND 2.5 MICROMETERS AVERAGE LENGTH; THEY
CC CONSIST OF ONLY A SINGLE POLYPEPTIDE CHAIN ARRANGED IN A HELICAL
CC CONFIGURATION OF FIVE SUBUNITS PER TURN IN THE ASSEMBLED PILUS.
CC -1- SIMILARITY: BELONGS TO THE N-ME-PHE PILI FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M37473; AAA23338.1; -.
DR PIR; PS0420; PS0420.
DR HSSP; P02974; 2PIL.
DR InterPro; IPR001082; Piliin.
DR InterPro; IPR001120; Prok_N_methyltn.
DR Pfam; PF00114; piliin; 1.
DR ProDom; PD000666; Piliin; 1.
DR PROSITE; PS00409; PROKAR_NTER_METHYL; 1.
KW Fimbria; Methylation.
FT CHAIN 1 160 FIMBRIAL PROTEIN.
FT MOD_RES 8 8 METHYLATION (BY SIMILARITY).
SQ SEQUENCE 160 AA; 16693 MW; 55ABC905E5B5057C CRC64;
Query Match 100.0%; Score 23; DB 1; Length 160;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DYDA 4
Db ||||
91 DYDA 94
RESULT 28
FMA7_BACNO STANDARD; PRT; 160 AA.
AC P27691;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 01-AUG-1992 (Rel. 23, Last annotation update)
DE Fimbrial protein precursor (Piliin) (Serogroup BI/AC127).
CN FINA.
OS Bacteroides nodosus (Dichelobacter nodosus).
OC Bacteria; Proteobacteria; gamma subdivision; Cardiobacteriaceae;
OX Dichelobacter.
OX NCBI_TaxID=870;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Serogroup B1 isolate AC127;
RX MEDLINE=91216435; PubMed=1673665;
RA Billington S.J., Rood J.I.;
RT "Sequence of fibrial subunit-encoding genes from virulent and benign
RL isolates of Dichelobacter (Bacteroides) nodosus.";
RL Gene 99:115-119(1991).
CC -1- SUBUNIT: THE PILI ARE POLAR FLEXIBLE FILAMENTS OF ABOUT 5.4
CC NANOMETERS DIAMETER AND 2.5 MICROMETERS AVERAGE LENGTH; THEY
CC CONSIST OF ONLY A SINGLE POLYPEPTIDE CHAIN ARRANGED IN A HELICAL
CC CONFIGURATION OF FIVE SUBUNITS PER TURN IN THE ASSEMBLED PILUS.

```

```

CC -1- SIMILARITY: BELONGS TO THE N-ME-PHE PILI FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M37472; AAA233340.1; -.
DR PIR; PS0419; PS0419.
DR HSSP; P02974; 1AY2.
DR InterPro; IPR001082; Piliin.
DR InterPro; IPR001120; Prok_N_methyltn.
DR Pfam; PF00114; piliin; 1.
DR ProDom; PD000666; Piliin; 1.
DR PROSITE; PS00409; PROKAR_NTER_METHYL; 1.
KW Fimbria; Methylation.
FT CHAIN 1 160 FIMBRIAL PROTEIN.
FT MOD_RES 8 8 METHYLATION (BY SIMILARITY).
SQ SEQUENCE 160 AA; 16674 MW; 7CBC909FF48287AD CRC64;
Query Match 100.0%; Score 23; DB 1; Length 160;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DYDA 4
Db ||||
91 DYDA 94
RESULT 29
GREA_MYCPN STANDARD; PRT; 160 AA.
ID P78019;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Transcription elongation factor greA (Transcript cleavage factor
DE greA).
DE GREA OR MPN401 OR MP437.
OS Mycoplasma pneumoniae.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2104;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 29342 / M129;
RX MEDLINE=97105885; PubMed=8948633;
RA Himmelreich R., Hilbert H., Plagens H., Pirkl E., Li B.-C.,
RA Herrmann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
RL pneumoniae.";
RL Nucleic Acids Res. 24:4420-4449(1996).
CC -1- FUNCTION: NECESSARY FOR EFFICIENT RNA POLYMERASE TRANSCRIPTION
CC ELONGATION PAST TEMPLATE-ENCODED ARRESTING SITES. THE ARRESTING
CC SITES IN DNA HAVE THE PROPERTY OF TRAPPING A CERTAIN FRACTION
CC OF ELONGATING RNA POLYMERASES THAT PASS THROUGH, RESULTING IN
CC LOCKED TERNARY COMPLEXES. CLEAVAGE OF THE NASCENT TRANSCRIPT BY
CC CLEAVAGE FACTORS SUCH AS GREA OR GREB ALLOWS THE RESUMPTION OF
CC ELONGATION FROM THE NEW 3' TERMINUS. GREA RELEASES SEQUENCES OF
CC 2 TO 3 NUCLEOTIDES (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE GREA/GREB FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----

```



```

DR EMBL; AE000043; AAB96085.1; -.
DR HSP; P21346; IGRJ.
DR InterPro; IPR001437; GreA_GreB.
DR Pfam; PF01272; GreA_GreB; 1.
DR Pfam; PF03449; GreA_GreB_N; 1.
DR ProDom; PD004918; GreA_GreB; 1.
DR PROSITE; PS00829; GREAB_1; 1.
DR PROSITE; PS00830; GREAB_2; 1.
KW Transcription regulation; DNA-binding; Coiled coil; Complete proteome.
FT DOMAIN 8 28 COILED COIL (POTENTIAL).
FT DOMAIN 48 75 COILED COIL (POTENTIAL).
SQ SEQUENCE 160 AA; 10101 MW; 32448C91712A19C2 CRC64;

Query Match
Best Local Similarity 100.0%; Score 23; DB 1; Length 160;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
DB 49 DYDA 52

RESULT 30
YEMO_YEAST STANDARD; PRT; 160 AA.
AC P40019;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-FEB-1995 (Rel. 31, Last annotation update)
DE Hypothetical 18.3 kDa protein in GAL83-YPT8 intergenic region.
GN YER030W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S28C / AB972;
RA Dietrich F.S., Mulligan J.T., Hennessey K.M., Allen E., Araujo R.,
RA Aviles E., Berno A., Brennan T., Carpenter J., Chen E., Cherry J.M.,
RA Chung E., Duncan M., Guzman E., Hartzell G., Hunkle-Smith S.,
RA Hymen R., Kayser A., Komp C., Lashkari D., Lew H., Lin D.,
RA Mesedale D., Nakahara K., Namath A., Norgren R., Oefner P., Oh C.,
RA Petel F.X., Roberts D., Sehl P., Schramm S., Shogren T., Smith V.,
RA Taylor P., Wei Y., Yelton M., Botstein D., Davis R.W.
RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U18778; AAB64563.1; -.
DR SGD; S0000832; YER030W.
KW Hypothetical protein
SQ SEQUENCE 160 AA; 19306 MW; 86BBE29FFE3D948B CRC64;

Query Match
Best Local Similarity 100.0%; Score 23; DB 1; Length 160;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
DB 47 DYDA 50

RESULT 31
FMA3_BACNO STANDARD; PRT; 161 AA.
ID FMA3_BACNO
AC P27689;

```

```

DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Fimbrial protein precursor (Pilin) (Serogroup B1/AC293).
GN FIMA.
OS Bacteroides nodosus (Dichelobacter nodosus).
OC Bacteria; Proteobacteria; gamma subdivision; Cardiobacteriaceae;
OC Dichelobacter.
OX NCBI_TaxID=870;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Serogroup B1 isolate AC293;
RC MEDLINE=91216435; PubMed=1673665;
RA Billington S.J., Rood J.I.;
RT "Sequence of fimbrial subunit-encoding genes from virulent and benign
RT isolates of Dichelobacter (Bacteroides) nodosus.";
RL Gene 99:115-119(1991).
CC
CC -1- SUBUNIT: THE PILI ARE POLAR FLEXIBLE FILAMENTS OF ABOUT 5.4
CC NANOMETERS DIAMETER AND 2.5 MICROMETERS AVERAGE LENGTH; THEY
CC CONSIST OF ONLY A SINGLE POLYPEPTIDE CHAIN ARRANGED IN A HELICAL
CC CONFIGURATION OF FIVE SUBUNITS PER TURN IN THE ASSEMBLED PILUS.
CC
CC -1- SIMILARITY: BELONGS TO THE N-ME-PHE PILI FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M37474; AAA23339.1; -.
DR PIR; PS0421; PS0421.
DR HSP; P02974; 2PIL.
DR InterPro; IPR001082; Pilin.
DR InterPro; IPR011120; Prok_N_methyltn.
DR Pfam; PF00114; pilin; 1.
DR ProDom; PD000666; Pilin; 1.
DR PROSITE; PS00409; PROKAR_NTER_METHYL; 1.
KW Fimbria; Methylation.
FT PROPEP 1 161 FIMBRIAL PROTEIN.
FT CHAIN 8 8 METHYLATION (BY SIMILARITY).
FT MOD_RES 8 8
SQ SEQUENCE 161 AA; 16860 MW; 3F0CF00DBF9EA385 CRC64;

Query Match
Best Local Similarity 100.0%; Score 23; DB 1; Length 161;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
DB 91 DYDA 94

RESULT 32
GREM_MYCGE STANDARD; PRT; 161 AA.
ID GREM_MYCGE
AC P47524;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Transcription elongation factor greA (Transcript cleavage factor
DE greA).
GN GREM OR MG282.
OS Mycoplasma genitalium.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2097;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 33530 / G-37; PubMed=7569993;
RC MEDLINE=96026346; PubMed=7569993;
RA Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,
RA Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,

```

```
RA Fritchman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhrmann J.L.,
RA Nguyen D.T., Doughterback T.R., Saudek D.M., Phillips C.A., Merrick J.M.,
RA Tomb J.F., Dougherty B.A., Bott K.F., Hu P.-C., Lucier T.S.,
RA Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;
RT "The minimal gene complement of Mycoplasma genitalium.";
RL Science 270:397-403(1995).
CC -!- FUNCTION: NECESSARY FOR EFFICIENT RNA POLYMERASE TRANSCRIPTION
CC ELONGATION PAST TEMPLATE-ENCODED ARRESTING SITES. THE ARRESTING
CC SITES IN DNA HAVE THE PROPERTY OF TRAPPING A CERTAIN FRACTION
CC OF ELONGATING RNA POLYMERASES THAT PASS THROUGH, RESULTING IN
CC LOCKED TERNARY COMPLEXES. CLEAVAGE OF THE NASCENT TRANSCRIPT BY
CC CLEAVAGE FACTORS SUCH AS GRE4 OR GREB ALLOWS THE RESUMPTION OF
CC ELONGATION FROM THE NEW 3' TERMINUS. GRE4 RELEASES SEQUENCES OF
CC 2 TO 3 NUCLEOTIDES (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE GRE4/GREB FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U39708; AAC71504.1; -.
DR HSSP: P21346; IGRJ.
DR TIGR: MG282; -.
DR InterPro: IPR001437; GreA_GreB.
DR Pfam: PF01272; GreA_GreB; 1.
DR Pfam: PF03449; GreA_GreB.N; 1.
DR ProDom: PD004918; GreA_GreB; 1.
DR PROSITE: PS00829; GREAB_1; 1.
DR PROSITE: PS00830; GREAB_2; 1.
DR Transcription regulation; DNA-binding; Coiled coil; Complete proteome.
KW DOMAIN 9 28 COILED COIL (POTENTIAL).
FT DOMAIN 47 76 COILED COIL (POTENTIAL).
FT SEQUENCE 161 AA; 18162 MW; 67E2850CB59BCC5B CRC64;
SQ
Query Match 100.0%; Score 23; DB 1; Length 161;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DYDA 4
DB 49 DYDA 52
TPX_STRGC STANDARD; PRT; 162 AA.
AC P42366;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Probable thiol peroxidase (EC 1.1.1.1.-).
GN TPX.
OS Streptococcus gordonii Challis.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OX NCBI_TaxID=29390;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PK488;
RX MEDLINE=95012638; PubMed=7927711;
RT Kolenbrander P.E., Andersen R.N., Ganeshkumar N.;
RT "Nucleotide sequence of the Streptococcus gordonii PK488
RT coaggregation adhesin gene, scaA, and ATP-binding cassette.";
RL Infect. Immun. 62:4469-4480(1994).
CC -!- FUNCTION: HAS ANTIOXIDANT ACTIVITY. COULD REMOVE PEROXIDES OR
CC H2O(2) (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE AHPC/TSA FAMILY. TPX SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M63481; AAC98427.1; -.
DR PIR: B43583; B43583.
DR InterPro: IPR000866; Ahpc-TSA.
DR Pfam: PF00578; Ahpc-TSA; 1.
DR PROSITE: PS01265; TPX; 1.
DR Oxidoreductase; Peroxidase.
KW INIT_MET 0
FT INIT_MET 0
SQ SEQUENCE 162 AA; 17754 MW; B48F965E238C7484 CRC64;
Query Match 100.0%; Score 23; DB 1; Length 162;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DYDA 4
DB 151 DYDA 154
TPX_STRSA STANDARD; PRT; 162 AA.
AC P31308;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable thiol peroxidase (EC 1.1.1.1.-).
GN TPX.
OS Streptococcus sanguis.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OX NCBI_TaxID=1305;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-40.
RC STRAIN=12;
RX MEDLINE=91147187; PubMed=1671775;
RT Ganeshkumar N., Hannam P.M., Kolenbrander P.E., McBride B.C.;
RT "Nucleotide sequence of a gene coding for a saliva-binding protein
RT (SsAb) from Streptococcus sanguis 12 and possible role of the protein
RT in coaggregation with actinomyces.";
RL Infect. Immun. 59:1093-1099(1991).
CC -!- FUNCTION: HAS ANTIOXIDANT ACTIVITY. COULD REMOVE PEROXIDES OR
CC H2O(2) (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE AHPC/TSA FAMILY. TPX SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L11577; AAA71948.1; -.
DR InterPro: IPR000866; Ahpc-TSA.
DR Pfam: PF002065; TPX.
DR Pfam: PF00578; Ahpc-TSA; 1.
DR PROSITE: PS01265; TPX; 1.
DR Oxidoreductase; Peroxidase.
KW INIT_MET 0
FT INIT_MET 0
SQ SEQUENCE 162 AA; 17781 MW; B49E665E238C7484 CRC64;
Query Match 100.0%; Score 23; DB 1; Length 162;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DYDA 4
DB 151 DYDA 154
```

```
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L11577; AAA71948.1; -.
DR InterPro: IPR000866; Ahpc-TSA.
DR Pfam: PF002065; TPX.
DR Pfam: PF00578; Ahpc-TSA; 1.
DR PROSITE: PS01265; TPX; 1.
DR Oxidoreductase; Peroxidase.
KW INIT_MET 0
FT INIT_MET 0
SQ SEQUENCE 162 AA; 17781 MW; B49E665E238C7484 CRC64;
Query Match 100.0%; Score 23; DB 1; Length 162;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DYDA 4
DB 151 DYDA 154
TPX_STRSA STANDARD; PRT; 162 AA.
AC P31308;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable thiol peroxidase (EC 1.1.1.1.-).
GN TPX.
OS Streptococcus sanguis.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OX NCBI_TaxID=1305;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-40.
RC STRAIN=12;
RX MEDLINE=91147187; PubMed=1671775;
RT Ganeshkumar N., Hannam P.M., Kolenbrander P.E., McBride B.C.;
RT "Nucleotide sequence of a gene coding for a saliva-binding protein
RT (SsAb) from Streptococcus sanguis 12 and possible role of the protein
RT in coaggregation with actinomyces.";
RL Infect. Immun. 59:1093-1099(1991).
CC -!- FUNCTION: HAS ANTIOXIDANT ACTIVITY. COULD REMOVE PEROXIDES OR
CC H2O(2) (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE AHPC/TSA FAMILY. TPX SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M63481; AAC98427.1; -.
DR PIR: B43583; B43583.
DR InterPro: IPR000866; Ahpc-TSA.
DR Pfam: PF00578; Ahpc-TSA; 1.
DR PROSITE: PS01265; TPX; 1.
DR Oxidoreductase; Peroxidase.
KW INIT_MET 0
FT INIT_MET 0
SQ SEQUENCE 162 AA; 17754 MW; B48F965E238C7484 CRC64;
Query Match 100.0%; Score 23; DB 1; Length 162;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DYDA 4
DB 151 DYDA 154
```



```

RESULT 37
YI75_STAAM
ID YI75_STAAM STANDARD; PRT; 171 AA.
AC Q53719;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein SAV1875/SA1692/MW1815 (ORF1).
GN SAV1875 OR SA1692 OR MW1815.
OS Staphylococcus aureus (strain Mu50 / ATCC 700699),
OS Staphylococcus aureus (strain N315),
OS Staphylococcus aureus (strain MW2), and
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=158878, 158879, 196620, 1280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Mu50 / ATCC 700699, and N315;
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-O., Ito T.,
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekinizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Hayashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus."
RL Lancet 357:1225-1240(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Mu50;
RX MEDLINE=22040717; PubMed=12044378;
RA Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,
RA Nagai Y., Iwama N., Asano K., Naimi T., Kuroda H., Cui L.,
RA Yamamoto K., Hiramatsu K.;
RT "Genome and virulence determinants of high virulence community-
RT acquired MRSA."
RL Lancet 359:1819-1827(2002).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=NCTC 8325;
RX MEDLINE=94131293; PubMed=7905453;
RA Borchardt S.A., Babwah A.V., Jayaswal R.K.;
RT "Sequence analysis of the region downstream from a peptidoglycan
RT hydrolase-encoding gene from Staphylococcus aureus NCTC8325."
RL Gene 137:253-258(1993).
CC -!- SIMILARITY: BELONGS TO THE THJ / PFPI FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; AF003363; BAB58037.1; -
DR EMBL; AP003135; BAB42961.1; -
DR EMBL; AP004828; BAB95680.1; -
DR EMBL; L19300; AAA18514.1; -
DR MEROPS; C56.0UP; -
DR InterPro; IPR002818; ThjJ.
DR Pfam; PF01965; ThjJ; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 171 AA; 18632 MW; A571A08B0FD0D719 CRC64;
Query Match 100.0%; Score 23; DB 1; Length 171;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 38
OLFA_RANPI
ID OLFA_RANPI STANDARD; PRT; 177 AA.
AC P06910;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Olfactory protein precursor.
DE Olfactory protein precursor.
OS Rana pipiens (Northern leopard frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranioidea; Ranidae; Rana.
OX NCBI_TaxID=8404;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Bowman's gland;
RX MEDLINE=87149013; PubMed=3493528;
RA Lee K.-H., Wells R.G., Reed R.R.;
RT "Isolation of an olfactory cDNA: similarity to retinol-binding
RT protein suggests a role in olfaction."
RL Science 235:1053-1056(1987).
CC -!- TISSUE SPECIFICITY: SYNTHESIZED IN THE BOWMAN'S GLANDS, WHICH
CC SECRETE THE MUCUS THAT BATHES THE CILIA OF THE OLFACTORY
CC NEUROEPITHELIUM.
CC -!- SIMILARITY: BELONGS TO THE LIPOCALIN FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; M15531; AAA49529.1; -
DR PIR; A25837; OVFGP.
DR InterPro; IPR000586; Lipocalin_cytfabp.
DR Pfam; PF00061; lipocalin; 1.
DR PROSITE; PS00213; LIPOCALIN; 1.
KW Olfaction; Signal; Lipocalin.
FT SIGNAL 1 13
FT CHAIN 14 177 OLFACTORY PROTEIN.
FT DISULFID 81 174 BY SIMILARITY.
SQ SEQUENCE 177 AA; 20191 MW; 61499AF22E72D251 CRC64;
Query Match 100.0%; Score 23; DB 1; Length 177;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
DB 64 DYDA 67

RESULT 39
SPI2_SOLTU
ID SPI2_SOLTU STANDARD; PRT; 186 AA.
AC P58515;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Serine protease inhibitor 2 (SPI-21) (SPI-21-5.2).
OS Solanum tuberosum (Potato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4113;
RN [1]
RP SEQUENCE, AND MASS SPECTROMETRY.
-----
```

RC STRAIN=cv. Istrinskii;
RX MEDLINE=21077058; PubMed=11209756;
RA Valueva T.A., Revina T.A., Mosolov V.V., Mentele R.;
RT "Primary structure of potato kunitz-type serine proteinase
inhibitor.";
RL Biol. Chem. 381:1215-1221(2000).
CC -!- FUNCTION: POTENT INHIBITOR OF SERINE PROTEASES (CHYMOTRYPSIN AND
TRYPSIN). INHIBITS TIGHTLY HUMAN LEUKOCYTE ELASTASE (HLE). DOES
NOT INHIBIT PAPAIN, PEPsin NOR CATHEPSIN D (CYSTEINE AND ASPARTIC
PROTEASES). PROTECTS THE PLANT BY INHIBITING PROTEASES OF INVADING
ORGANISMS, DECREASING BOTH HYPHAL GROWTH AND ZOOSPORES GERMINATION
OF PHYTOPHTHORA INFESTANS.
CC -!- SUBUNIT: HETERODIMER OF CHAINS A AND B; DISULFIDE-LINKED. DOUBLE-
HEADED INHIBITOR ABLE TO FORM TRIPLE COMPLEXES WITH TARGET
PROTEASES, BY BINDING SIMULTANEOUSLY ONE MOLECULE OF TRYPSIN AND
ONE MOLECULE OF CHYMOTRYPSIN.
CC -!- SUBCELLULAR LOCATION: Vacuolar (By similarity).
CC -!- TISSUE SPECIFICITY: TUBERS.
CC -!- INDUCTION: By infection with Phytophthora infestans.
CC -!- PTM: PROBABLY SYNTHESIZED AS A SINGLE-CHAIN PRECURSOR WHICH
IS CLEAVED TO FORM CHAIN A AND CHAIN B.
CC -!- MASS SPECTROMETRY: MW=16396; MW_ERR=16; METHOD=Electrospray;
RANGE=1-150.
CC -!- MASS SPECTROMETRY: MW=4182; MW_ERR=4; METHOD=Electrospray;
RANGE=151-187.
CC -!- MISCELLANEOUS: THE SERINE PROTEASE INHIBITOR PSPI-21 COMPRISES TWO
PROTEIN SPECIES WITH PI 5.2 AND 6.3, DENOTED AS PSPI-21-5.2 AND
PSPI-21-6.3. THEY MAY BE ENCODED BY TWO ALLELES AT THE SAME GENE
LOCUS.
CC -!- SIMILARITY: BELONGS TO THE LEGUMINOUS KUNITZ-TYPE INHIBITOR
FAMILY.
DR InterPro: IPR002160; Kunitz_legume.
DR ProDom: PD000891; Kunitz_legume; 1.
DR SMART: SM00452; STI: 1.
DR PROSITE: PS00283; SOYBEAN_KUNITZ; 1.
KW Serine protease inhibitor; Multigene family.
FT CHAIN 1 150 CHAIN A.
FT NON_CONS 150 151
FT CHAIN 151 186 CHAIN B.
FT ACT_SITE 67 68 REACTIVE BOND (FOR TRYPSIN).
FT ACT_SITE 115 116 REACTIVE BOND (FOR CHYMOTRYPSIN OR HLE).
FT DISULFID 48 97 BY SIMILARITY.
FT DISULFID 146 157 INTERCHAIN (BY SIMILARITY).
SQ SEQUENCE 186 AA; 20115 MW; 26C19E0BB8A01E26 CRC64;
Query Match 100.0%; Score 23; DB 1; Length 186;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DYDA 4
DB 107 DYDA 110
RESULT 40
YZZE_STRCO STANDARD; PRT; 186 AA.
AC Q9K470;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein SC07215.
GN SC07215 OR SC2H12.14C.
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomycetes.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RX MEDLINE=2196410; PubMed=12000953;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,

RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,
RA Rabinowitz E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorek A., Woodward J., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomycetes
coelicolor A3(2).";
RL Nature 417:141-147(2002).
CC -!- SIMILARITY: BELONGS TO THE UPF0157 (GRPB) FAMILY.
CC -!- THIS SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AL359215; CAB94638.1; .
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 186 AA; 20737 MW; 38545B835F9C09DB CRC64;
Query Match 100.0%; Score 23; DB 1; Length 186;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DYDA 4
DB 30 DYDA 33
RESULT 41
SPII_SOLTU STANDARD; PRT; 187 AA.
AC P58514;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Serine protease inhibitor 1 (PSPI-21) (PSPI-21-6.3).
OS Solanum tuberosum (Potato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4113;
RN [1]
RP SEQUENCE, MASS SPECTROMETRY, AND DISULFIDE BONDS.
RC STRAIN=cv. Istrinskii;
RX MEDLINE=21077058; PubMed=11209756;
RA Valueva T.A., Revina T.A., Mosolov V.V., Mentele R.;
RT "Primary structure of potato kunitz-type serine proteinase
inhibitor.";
RL Biol. Chem. 381:1215-1221(2000).
RN [2]
RP SEQUENCE OF 1-25 AND 151-160.
RC STRAIN=cv. Istrinskii;
RX MEDLINE=98258940; PubMed=9598993;
RA Valueva T.A., Revina T.A., Kladnitskaya G.V., Mosolov V.V.;
RT "Kunitz-type proteinase inhibitors from intact and
Phytophthora-infected potato tubers.";
RL FEBS Lett. 426:131-134(1998).
CC -!- FUNCTION: POTENT INHIBITOR OF SERINE PROTEASES (CHYMOTRYPSIN AND
TRYPSIN). INHIBITS TIGHTLY HUMAN LEUKOCYTE ELASTASE (HLE). DOES
NOT INHIBIT PAPAIN, PEPsin NOR CATHEPSIN D (CYSTEINE AND ASPARTIC
PROTEASES). PROTECTS THE PLANT BY INHIBITING PROTEASES OF INVADING
ORGANISMS, DECREASING BOTH HYPHAL GROWTH AND ZOOSPORES GERMINATION
OF PHYTOPHTHORA INFESTANS.
CC -!- SUBUNIT: HETERODIMER OF CHAINS A AND B; DISULFIDE-LINKED. DOUBLE-
HEADED INHIBITOR ABLE TO FORM TRIPLE COMPLEXES WITH TARGET
PROTEASES, BY BINDING SIMULTANEOUSLY ONE MOLECULE OF TRYPSIN AND
ONE MOLECULE OF CHYMOTRYPSIN.
CC -!- SUBCELLULAR LOCATION: Vacuolar (By similarity).

```

CC CC -!- TISSUE SPECIFICITY: TUBERS.
CC CC -!- INDUCTION: By infection with Phytophthora infestans.
CC CC -!- PTM: PROBABLY SYNTHESIZED AS A SINGLE-CHAIN PRECURSOR WHICH
CC CC IS CLEAVED TO FORM CHAIN A AND CHAIN B.
CC CC -!- MASS SPECTROMETRY: MW=16078; MW_ERR=16; METHOD=Electrospray;
CC CC RANGE=1-150.
CC CC -!- MASS SPECTROMETRY: MW=4282; MW_ERR=4; METHOD=Electrospray;
CC CC RANGE=151-187.
CC CC -!- MISCELLANEOUS: THE SERINE PROTEASE INHIBITOR PSPI-21 COMPRISES TWO
CC CC PROTEIN SPECIES WITH PI 5.2 AND 6.3, DENOTED AS PSPI-21-5.2 AND
CC CC PSPI-21-6.3. THEY MAY BE ENCODED BY TWO ALLELES AT THE SAME GENE
CC CC LOCUS.
CC CC -!- SIMILARITY: BELONGS TO THE LEGUMINOUS KUNITZ-TYPE INHIBITOR
CC CC FAMILY.
CC DR InterPro; IPR002160; Kunitz_legume.
CC DR ProDom; PD000891; Kunitz_legume; 1.
CC DR SMART; SM00452; ST1; 1.
CC DR PROSITE; PS00283; SOYBEAN_KUNITZ; 1.
CC KW Serine protease inhibitor; Multigene family.
CC FT CHAIN 1 150 CHAIN A.
CC FT NON_CONS 150 151
CC FT CHAIN 151 187 CHAIN B.
CC FT ACT_SITE 67 68 REACTIVE BOND (FOR TRYPSIN).
CC FT ACT_SITE 115 116 REACTIVE BOND (FOR CHYMOTRYPSIN OR HLE).
CC FT DISULFID 48 97
CC FT DISULFID 146 157 INTERCHAIN.
CC SQ SEQUENCE 187 AA; 20243 MW; 6AFE7F4AAAFD706B CRC64;

Query Match 100.0%; Score 23; DB 1; Length 187;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
DB 107 DYDA 110

RESULT 42
WRBA_ECOLI STANDARD; PRT; 197 AA.
AC P30849; P75890; P77543;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Flavoprotein wrba (Trp repressor binding protein).
GN WRBA OR B1004.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
[1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-18.
RC STRAIN-K12;
RX MEDLINE=93296226; PubMed=8516330;
RA Yang W., Ni L., Somerville R.L.;
RT "A stationary-phase protein of Escherichia coli that affects the mode
RT of association between the trp repressor protein and operator-bearing
RT DNA".
RL Proc. Natl. Acad. Sci. U.S.A. 90:5796-5800(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Bernal N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-K12;
RX MEDLINE=97061202; PubMed=8905232;

```

```

RA Oshima T., Aiba H., Baba T., Fujita K., Hayashi K., Honjo A.,
RA Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,
RA Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,
RA Mori H., Motomura K., Nakamura Y., Nishimoto H., Nishio Y., Saito N.,
RA Sampei G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,
RA Yano M., Horiuchi T.;
RT "A 718-kb DNA sequence of the Escherichia coli K-12 genome
RT corresponding to the 12.7-28.0 min region on the linkage map.";
RL DNA Res. 3:137-155(1996).
RN [4]
RP SEQUENCE OF 1-10, AND CHARACTERIZATION.
RC STRAIN-JM101;
RX MEDLINE=98361966; PubMed=9694845;
RA Grandori R., Khalifah P., Boice J.A., Fairman R., Giovanielli K.,
RA Carey J.;
RT "Biochemical characterization of wrba, founding member of a new family
RT of multimeric flavodoxin-like proteins.";
RL J. Biol. Chem. 273:20960-20966(1998).
CC -!- FUNCTION: Seems to enhance the formation and/or stability of
CC noncovalent complexes between the trp repressor protein and
CC operator-bearing DNA. However, wrba does not specifically
CC influence the affinity or mode of binding of trpR to its operator.
CC It seems possible that the association may have a structural,
CC rather than functional, basis. Wrba alone does not interact with
CC the operator-bearing DNA; the trpR/trpO complex is probably the
CC major target species. The wrba protein may function as an
CC accessory element in blocking trpR-specific transcriptional
CC processes that might be physiologically disadvantageous in the
CC stationary phase of the bacterial life cycle.
CC -!- COFACTOR: Binds 1 FMN per monomer.
CC -!- SUBUNIT: Homodimer and homotetramer; in equilibrium.
CC -!- SIMILARITY: BELONGS TO THE WRBA FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M99166; AAA24759.1; -
CC EMBL; AE000202; AAC74089.1; -
CC EMBL; D90738; BAA35781.1; -
CC EMBL; D90737; BAA35771.1; -
CC SWISS-2DPAGE; P30849; COLI.
CC EcoGene; EG11540; wrba.
CC InterPro; IPR001226; Flavodoxin.
CC Fram; PF00258; flavodoxin; 1.
CC DR PROSITE; PS00201; FLAVODOXIN; FALSE_NEG.
CC KW Flavoprotein; FMN; Complete proteome.
CC FT INIT_MET 0
CC FT CONFLICT 141 141 G -> A (IN REF. 1).
CC SQ SEQUENCE 197 AA; 20714 MW; 7B48C7CC2B71747 CRC64;

Query Match 100.0%; Score 23; DB 1; Length 197;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
DB 67 DYDA 70

RESULT 43
WRBA_SALTI STANDARD; PRT; 197 AA.
AC Q827N9;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Flavoprotein wrba (Trp repressor binding protein).
GN WRBA OR STY1155.

```

```

OS Salmonella typhi.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBI_TaxID=601;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=CT18;
RX MEDLINE=21534947; PubMed=11677608;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahia M.,
RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA Felwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,
RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrell B.G.;
RT "Complete genome sequence of a multiple drug resistant Salmonella
RT enterica serovar Typhi CT18."
RL Nature 413:848-852(2001).
CC -!- FUNCTION: Seems to enhance the formation and/or stability of
CC noncovalent complexes between the trp repressor protein and
CC operator-bearing DNA (By similarity).
CC -!- COFACTOR: Binds 1 FMN per monomer (By similarity).
CC -!- SIMILARITY: BELONGS TO THE WRBA FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AL627269; CAD08244.1; -
CC InterPro: IPR001226; Flavodoxin.
CC Pfam: PF00258; flavodoxin; 1.
CC PROSITE: PS00201; FLAVODOXIN; FALSE_NEG.
KW Flavoprotein; FMN; Complete proteome.
FT INIT_MET 0
FT BY SIMILARITY.
SQ SEQUENCE 197 AA; 20706 MW; DCD65E43E98C7112 CRC64;

Query Match 100.0%; Score 23; DB 1; Length 197;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
DB 67 DYDA 70

RESULT 44
WRBA_SALTY
ID WRBA_SALTY STANDARD; PRT; 197 AA.
AC Q82Q40;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Flavoprotein wrba (Trp repressor binding protein).
GN WRBA OR STM119.
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBI_TaxID=602;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium

```

```

RT Nature 413:852-856(2001).
RL -!- FUNCTION: Seems to enhance the formation and/or stability of
RL noncovalent complexes between the trp repressor protein and
RL operator-bearing DNA (By similarity).
CC -!- COFACTOR: Binds 1 FMN per monomer (By similarity).
CC -!- SIMILARITY: BELONGS TO THE WRBA FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE008748; AAL20051.1; -
CC StyGene; SG7777; wrba.
CC InterPro: IPR001226; Flavodoxin.
CC Pfam: PF00258; Flavodoxin; 1.
CC PROSITE: PS00201; FLAVODOXIN; FALSE_NEG.
KW Flavoprotein; FMN; Complete proteome.
FT INIT_MET 0
FT BY SIMILARITY.
SQ SEQUENCE 197 AA; 20736 MW; C52F61B136A185C6 CRC64;

Query Match 100.0%; Score 23; DB 1; Length 197;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
DB 67 DYDA 70

RESULT 45
SODI_HALNI
ID SODI_HALNI STANDARD; PRT; 199 AA.
AC P09737; Q9HQF1;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Superoxide dismutase [Mn] 1 (EC 1.15.1.1).
GN SOD1 OR SOD OR VNG1190G.
OS Halobacterium sp. (strain NRC-1), and
OS Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC Halobacteriaceae; Halobacterium.
OX NCBI_TaxID=64091, 2242;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=NRC-1;
RX MEDLINE=20504483; PubMed=11016950;
RA Ng W.V., Kennedy S.P., Mahairas G.G., Berquist B., Pan M.,
RA Shukla H.D., Lasky S.R., Baliga N.S., Thorsson V., Sbrogna J.,
RA Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A.,
RA Leithauser B., Keller K., Cruz R., Danson M.J., Hough D.W.,
RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,
RA Isenbarger T.A., Peck R.F., Pohlschroder M., Spudich J.L., Jung K.-H.,
RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
RA Ehardt H., Lowe T.M., Liang P., Riley M., Hood L., DasSarma S.;
RT "Genome sequence of Halobacterium species NRC-1."
RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
RN [2]
RP SEQUENCE FROM N.A.
RX SPECIES=H.cutirubrum;
RX MEDLINE=89308646; PubMed=2745441;
RA May B.P., Dennis P.;
RT "Evolution and regulation of the gene encoding superoxide dismutase
RT from the archaeobacterium Halobacterium cutirubrum."
RL J. Biol. Chem. 264:12253-12258(1989).
RN [3]
RP SEQUENCE FROM N.A.
RX SPECIES=H.cutirubrum;

```

```

RX MEDLINE=89248676; PubMed=2720491;
RA May B.P., Tam P., Dennis P.P.;
RT "The expression of the superoxide dismutase gene in Halobacterium
RL cutirubrum and Halobacterium volcanii."
RN Can. J. Microbiol. 35:171-175(1989).
RN [4]
RP SEQUENCE OF 1-56.
RC SPECIES=H.cutirubrum;
RX MEDLINE=87165744; PubMed=3104309;
RA May B.P., Dennis P.P.;
RT "Superoxide dismutase from the extremely halophilic archaeobacterium
RL Halobacterium cutirubrum."
RN J. Bacteriol. 169:1417-1422(1987).
CC -!- FUNCTION: Destroys radicals which are normally produced within the
CC cells and which are toxic to biological systems.
CC -!- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).
CC -!- COFACTOR: Manganese.
CC -!- SUBUNIT: Homotetramer (Probable).
CC -!- SIMILARITY: BELONGS TO THE IRON/MANGANESE SUPEROXIDE DISMUTASE
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE005046; AAG19564.1; -.
DR EMBL; J04956; AAA72217.1; -.
DR PIR; A26932; A26932.
DR PIR; A34319; A34319.
DR PIR; S11601; S11601.
DR HSP; P17670; 1IDS.
DR InterPro; IPR001189; SODismutase.
DR Pfam; PF00081; sodfe; 1.
DR Pfam; PF02777; sodfe_C; 1.
DR ProDom; PD000475; SODismutase; 1.
DR PROSITE; PS00088; SOD_MN; 1.
KW Oxidoreductase; Manganese; Complete proteome.
FT INIT_MET 0
FT METAL 28 28 MANGANESE (BY SIMILARITY).
FT METAL 75 75 MANGANESE (BY SIMILARITY).
FT METAL 157 157 MANGANESE (BY SIMILARITY).
FT METAL 161 161 MANGANESE (BY SIMILARITY).
FT CONFLICT 1 1 S -> T (IN REF. 4).
FT CONFLICT 51 51 E -> N (IN REF. 4).
SQ SEQUENCE 199 AA; 22254 MW; 2D3CC5664F811719 CRC64;

```

```

Query Match          100.0%; Score 23; DB 1; Length 199;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 DYDA 4
   |||
Db 11 DYDA 14

```

Search completed: February 6, 2003, 11:16:42
Job time : 8.33333 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OW protein - protein search, using sw model

Run on: February 6, 2003, 11:16:00 ; Search time 21.3333 Seconds
(without alignments)
38.634 Million cell updates/sec

Title: PAT943-1

Perfect score: 23

Sequence: 1 dyda 4

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 1257

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 100%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_21.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rvirus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	23	100.0	42	2 Q49040	Q49040 mycoplasma
2	23	100.0	45	17 Q82X27	Q82X27 pyrobaculum
3	23	100.0	52	5 O17545	O17545 caenorhabdi
4	23	100.0	58	10 Q94I94	Q94I94 atropa bell
5	23	100.0	63	17 Q8TV22	Q8TV22 methanopyru
6	23	100.0	67	16 Q92JX2	Q92JX2 rhizobium m
7	23	100.0	67	16 Q8U9H5	Q8U9H5 agrobacteri
8	23	100.0	68	16 Q8UJW5	Q8UJW5 agrobacteri
9	23	100.0	69	16 Q8VJF8	Q8VJF8 mycobacteri
10	23	100.0	70	4 Q96IE9	Q96IE9 homo sapien
11	23	100.0	72	10 Q9SWZ0	Q9SWZ0 lycopersico
12	23	100.0	79	2 P82571	P82571 streptococc
13	23	100.0	81	2 O70101	O70101 streptococc
14	23	100.0	85	2 O70065	O70065 streptococc
15	23	100.0	86	7 Q9TNR9	Q9TNR9 homo sapien
16	23	100.0	86	16 Q97LK6	Q97LK6 clostridium

17	23	100.0	91	16 Q9RXH2	Q9RXH2 deinococcus
18	23	100.0	92	12 Q72133	Q72133 human papil
19	23	100.0	92	12 Q72134	Q72134 human papil
20	23	100.0	92	12 Q72138	Q72138 human papil
21	23	100.0	92	12 Q72141	Q72141 human papil
22	23	100.0	93	16 Q8XE95	Q8XE95 escherichia
23	23	100.0	95	16 Q82I04	Q82I04 yersinia pe
24	23	100.0	96	13 Q9DPF6	Q9DPF6 oncorhynchu
25	23	100.0	100	16 Q07201	Q07201 mycobacteri
26	23	100.0	100	16 Q49990	Q49990 mycobacteri
27	23	100.0	102	5 Q23455	Q23455 caenorhabdi
28	23	100.0	103	9 Q48416	Q48416 bacterioph
29	23	100.0	104	2 Q30472	Q30472 bacillus su
30	23	100.0	105	2 Q9L9H2	Q9L9H2 enterococcu
31	23	100.0	105	16 Q34372	Q34372 bacillus su
32	23	100.0	107	16 Q8YI88	Q8YI88 brucella me
33	23	100.0	109	10 Q81229	Q81229 zea mays (m
34	23	100.0	109	16 Q9A859	Q9A859 caulobacter
35	23	100.0	109	17 Q9HNJ7	Q9HNJ7 halobacteri
36	23	100.0	110	12 Q8V6T9	Q8V6T9 halovirus h
37	23	100.0	111	5 Q9BIF5	Q9BIF5 manestra br
38	23	100.0	112	9 Q80205	Q80205 methanobact
39	23	100.0	112	16 Q9CBA7	Q9CBA7 mycobacteri
40	23	100.0	112	16 Q06425	Q06425 mycobacteri
41	23	100.0	113	1 Q9HH48	Q9HH48 methanobact
42	23	100.0	114	4 Q9H0N8	Q9H0N8 homo sapien
43	23	100.0	114	16 Q9AB13	Q9AB13 caulobacter
44	23	100.0	117	2 Q9EXU4	Q9EXU4 rhizobium m
45	23	100.0	117	11 Q8VHK4	Q8VHK4 mus musculu

ALIGNMENTS

RESULT 1

Q49040
ID Q49040 PRELIMINARY; PRT; 42 AA.
AC Q49040;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Transcription elongation factor (Fragment).
OS Mycoplasma capricolum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Entomoplasmatales; Entomoplasmataceae.
OX NCBI_TaxID=2095;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 27343(KID);
RX MEDLINE=96059641; PubMed=7476192;
RA Bork P., Ouzounis C., Casari G., Schneider R., Sander C., Dolan M.,
RA Gilbert W., Gillevet P.M.;
RT "Exploring the Mycoplasma capricolum genome: a minimal cell reveals
its physiology.";
RL Mol. Microbiol. 16:955-967(1995).
DR EMBL; Z33203; CAA83788.1;
DR InterPro; IPR001437; GreA_Greb.
DR Pfam; PF03449; GreA_GrebL_N; 1.
DR ProDom; PD004918; GreA_Greb; 1.
KW Elongation factor.
FT NON_TER 1
SQ SEQUENCE 42 AA; 4755 MW; 4A4A5D1D7DED6864 CRC64;

Query Match 100.0%; Score 23; DB 2; Length 42;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
DB 29 DYDA 32

RESULT 2

```
Q8ZXZ7
ID Q8ZXZ7 PRELIMINARY; PRT; 45 AA.
AC Q8ZXZ7;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DE 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Conserved within P. aerophilum.
GN PAE1015.
OS Pyrobaculum aerophilum.
OC Archaea; Crenarchaeota; Thermoprotei; Thermoproteales;
OC Thermoproteaceae; Pyrobaculum.
OX NCBI_TaxID=13773;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=IM2 / ATCC 51768 / DSM 7523;
RX PubMed=11792869;
RA Fitz-Gibbon S.T., Ladner H., Kim U.-J., Stetter K.O., Simon M.I.,
RA Miller J.H.;
RT "Genome sequence of the hyperthermophilic crenarchaeon Pyrobaculum
RT aerophilum.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:984-989(2002).
DR EMBL; AE009799; AAL63199.1; -.
KW Complete proteome.
SQ SEQUENCE 45 AA; 4568 MW; 8829F6FC79E4B5EC CRC64;

Query Match 100.0%; Score 23; DB 17; Length 45;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
   ||||
Db 30 DYDA 33

RESULT 3
O17545
ID O17545 PRELIMINARY; PRT; 52 AA.
AC O17545;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DE 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Protein tyrosine kinase (EC 2.7.1.112) (Fragment).
GN CEHD-28.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
SEQUENCE FROM N.A.
RA Oates A.C., Wilks A.F.;
RT "How many PKs to organise a worm?";
RL The Worm Breeders Gazette 14:87-87(1995).
RN [2]
SEQUENCE FROM N.A.
RA Oates A.C., Wollberg P., Achen M.G., Wilks A.F.;
RT "Sampling the Genomic Pool of Protein Tyrosine Kinase Genes using the
RT Polymerase Chain Reaction with Genomic DNA.";
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ002906; CAA05741.1; -.
DR InterPro; IPR00719; Euk_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Euk_pkinase; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
KW Kinase; Transferase.
FT NON_TER 1
FT NON_TER 52
SQ SEQUENCE 52 AA; 5781 MW; ACC807B59F33AB31 CRC64;

Query Match 100.0%; Score 23; DB 5; Length 52;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
```

```
Db 21 DYDA 24
   ||||
RESULT 4
Q94I94
ID Q94I94 PRELIMINARY; PRT; 58 AA.
AC Q94I94;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Putative ribosomal-like protein (Fragment).
DE Atropa belladonna (Belladonna) (Deadly nightshade).
OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Atropa.
OX NCBI_TaxID=33113;
RN [1]
SEQUENCE FROM N.A.
RA Nouar E., Baucher M., Jaziri M.;
RT "Differential gene expression in Atropa belladonna leafy gall.";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
SEQUENCE FROM N.A.
RA Nouar E.;
RL Thesis (2001), Department of Plant Biotechnology,
RL Universite Libre de Bruxelles, Brussels, Belgium.
DR EMBL; AJ309378; CAC40748.1; -.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 58 AA; 6778 MW; 4C097DAA58C3CBAD CRC64;

Query Match 100.0%; Score 23; DB 10; Length 58;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
   ||||
Db 53 DYDA 56

RESULT 5
Q8TV22
ID Q8TV22 PRELIMINARY; PRT; 63 AA.
AC Q8TV22;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Ferredoxin.
GN MK1579.
OS Methanopyrus kandleri.
OC Archaea; Euryarchaeota; Methanopyri; Methanopyrales; Methanopyraceae;
OC Methanopyrus.
OX NCBI_TaxID=2320;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=AV19 / DSM 6324 / JCM 9639;
RX MEDLINE=21927647; PubMed=11930014;
RA Slesarev A.I., Mezhevaya K.V., Makarova K.S., Polushin N.N.,
RA Shcherbinina O.V., Shakhova V.V., Belova G.I., Aravind L.,
RA Natile D.A., Rogozin I.B., Tatusov R.L., Wolf Y.I., Stetter K.O.,
RA Malykh A.G., Koonin E.V., Kozyavkin S.A.;
RT "The complete genome of hyperthermophile Methanopyrus kandleri AV19
RT and monophyly of archaeal methanogens.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:4644-4649(2002).
DR EMBL; AE010447; AAM02792.1; -.
KW Complete proteome.
SQ SEQUENCE 63 AA; 6690 MW; 55DDC3C8529EEC21 CRC64;

Query Match 100.0%; Score 23; DB 17; Length 63;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
```

```
Db 7 DYDA 10
||||
RX Q92JX2 PRELIMINARY; PRT; 67 AA.
AC Q92JX2;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical protein R02784.
GN R02784 OR SMC04000.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1021;
RX MEDLINE=21396507; PubMed=11481430;
RA Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,
RA Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,
RA Godrie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masuy D.,
RA Pohl T., Portetelle D., Puehler A., Purnelle B., Ramsperger U.,
RA Renard C., Thebaud P., Vandenbol M., Weidner S., Galibert F.;
RT "Analysis of the chromosome sequence of the legume symbiont
RT Sinorhizobium meliloti strain 1021.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).
DR EMBL; AL591791; CAC47363.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 67 AA; 7733 MW; 94416A9E2191AF32 CRC64;

Query Match 100.0%; Score 23; DB 16; Length 67;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
||||
Db 22 DYDA 25

RESULT 7
Q8U9H5 PRELIMINARY; PRT; 67 AA.
AC Q8U9H5;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hypothetical protein Atu3753.
GN ATU3753 OR AGR_L_2160.
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=176299;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608550; PubMed=11743193;
RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
RA Okura Y.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. Sr.,
RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
RA Kutayavin T., Levy R., Li M.-J., McClelland E., Palmieri A.,
RA Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon D.,
RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
RA Nester E.W.;
RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
RT C58.";
RL Science 294:2317-2323(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608551; PubMed=11743194;
RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
RA Qurollo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,
RA Houmlel K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu F.,
RA Wollam C., Allinger M., Doughty D., Scott C., Lappas C., Markelz B.,
RA Flanagan C., Crowell C., Gursun J., Lomo C., Sear C., Strub G.,
RA Cielo C., Slater S.;
RT "Genome sequence of the plant pathogen and biotechnology agent
RT Agrobacterium tumefaciens C58.";
RL Science 294:2323-2328(2001).
DR EMBL; AE008965; AAL46138.1; -.
DR EMBL; AE007913; AAK90826.1; -.
KW Hypothetical protein; Plasmid; Complete proteome.
SQ SEQUENCE 68 AA; 7596 MW; E42F36903AA88F86 CRC64;

Query Match 100.0%; Score 23; DB 16; Length 68;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
||||
Db 22 DYDA 25

RESULT 8
Q8UJM5 PRELIMINARY; PRT; 68 AA.
AC Q8UJM5;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hypothetical protein Atu5451.
GN ATU5451 OR AGR_PAT_562.
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=176299;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608550; PubMed=11743193;
RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
RA Okura Y.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. Sr.,
RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
RA Kutayavin T., Levy R., Li M.-J., McClelland E., Palmieri A.,
RA Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon D.,
RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
RA Nester E.W.;
RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
RT C58.";
RL Science 294:2317-2323(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608551; PubMed=11743194;
RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
RA Qurollo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,
RA Houmlel K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu F.,
RA Wollam C., Allinger M., Doughty D., Scott C., Lappas C., Markelz B.,
RA Flanagan C., Crowell C., Gursun J., Lomo C., Sear C., Strub G.,
RA Cielo C., Slater S.;
RT "Genome sequence of the plant pathogen and biotechnology agent
RT Agrobacterium tumefaciens C58.";
RL Science 294:2323-2328(2001).
DR EMBL; AE008965; AAL46138.1; -.
DR EMBL; AE007913; AAK90826.1; -.
KW Hypothetical protein; Plasmid; Complete proteome.
SQ SEQUENCE 68 AA; 7596 MW; E42F36903AA88F86 CRC64;
```

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
|||||
Db 33 DYDA 36

RESULT 9
Q8VJF8 ID Q8VJF8 PRELIMINARY; PRT; 69 AA.
AC Q8VJF8;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DE Hypothetical protein MT2627.1.
GN MT2627.1.
OS Mycobacterium tuberculosis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / OSHKOSH;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
laboratory strains."
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE007097; AKA46939.1; -.
DR TIGR; MT2627; -.
KW Hypothetical protein.
SQ SEQUENCE 69 AA; 7386 MW; 2F23B964568A2046 CRC64;

Query Match 100.0%; Score 23; DB 16; Length 69;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
|||||
Db 43 DYDA 46

RESULT 10
Q96IE9 ID Q96IE9 PRELIMINARY; PRT; 70 AA.
AC Q96IE9;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DE Unknown (protein for MGC:15572).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PLACENTA;
RA Strausberg R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC007583; AAH07583.1; -.
DR InterPro; IPR001014; Ribosomal_L23.
DR Pfam; PF00276; Ribosomal_L23; 1.
DR PRODOM; PD001141; Ribosomal_L23; 1.
DR PROSITE; PS00050; RIBOSOMAL_L23; UNKNOWN_1.
SQ SEQUENCE 70 AA; 7923 MW; AC1C466548F343C1 CRC64;

Query Match 100.0%; Score 23; DB 4; Length 70;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
|||||
Db 57 DYDA 60

RESULT 11
Q9SWZ0 ID Q9SWZ0 PRELIMINARY; PRT; 72 AA.
AC Q9SWZ0;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE Hypoxia-induced hypothetical protein 162 (Fragment).
OS Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4081;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. AILSA CRAIG;
RA Aggelis A., Kanellis A.K.;
RT "Hypoxia-induced cDNAs isolated by RNA differential display."
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF081024; AAD46143.1; -.
KW Hypothetical protein.
FT NON_TER 72 72
SQ SEQUENCE 72 AA; 7845 MW; 2C1B214D33B2184E CRC64;

Query Match 100.0%; Score 23; DB 10; Length 72;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
|||||
Db 68 DYDA 71

RESULT 12
P82571 ID P82571 PRELIMINARY; PRT; 79 AA.
AC P82571;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE Unknown protein from 2D-page (Fragments).
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
OC Streptococcaceae; Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE, AND MASS SPECTROMETRY.
RC STRAIN=JRS4;
RA Hogan D.A., Du P., Stevenson T.I., Whitton M., Kilby G.W., Rogers J.,
RA VanBogelen R.A.;
RT "Two-dimensional gel electrophoresis map of Streptococcus pyogenes
proteins."
RL Submitted (MAY-2000) to the SWISS-PROT data bank
CC -I- MASS SPECTROMETRY; MW=49603.57; METHOD=ELECTROSPRAY.
DR InterPro; IPR001327; FAD_pyr_redox.
DR PRINTS; PR001100; FADPNR.
DR PRINTS; PR00411; PNDRTASE1.
FT NON_TER 1 1
FT NON_CONS 15 16
FT NON_CONS 30 31
FT NON_CONS 47 48
FT NON_CONS 66 67
FT NON_TER 79 79
SQ SEQUENCE 79 AA; 8549 MW; 4ACF6EAD96A49F7A CRC64;

Query Match 100.0%; Score 23; DB 2; Length 79;
Best Local Similarity 100.0%; Pred. No. 3.6e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
||||
DB 11 DYDA 14

RESULT 13
070101 ID 070101 PRELIMINARY; PRT; 81 AA.
AC 070101;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-AUG-1998 (TrEMBLrel. 07, Last annotation update)
DE PNEUMOCOCCAL surface protein A (Fragment).
GN PSPA.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
OC Streptococcaceae; Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1196;
RA Robinson D.A., Turner J.S., Facklam R.R., Parkinson A.J.,
RA Breiman R.F., Gratten M., Hollingshead S.K., Briles D.E., Crain M.J.;
RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF036624; AAC09243.1; -;
DR EMBL; AF036621; AAC09240.1; -;
DR EMBL; AF036622; AAC09241.1; -;
DR EMBL; AF036623; AAC09242.1; -;
FT NON_TER 81
SQ SEQUENCE 81 AA; 8879 MW; 3752E31D97028C0A CRC64;

Query Match 100.0%; Score 23; DB 2; Length 81;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
||||
DB 45 DYDA 48

RESULT 14
070065 ID 070065 PRELIMINARY; PRT; 85 AA.
AC 070065;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-AUG-1998 (TrEMBLrel. 07, Last annotation update)
DE PNEUMOCOCCAL surface protein A (Fragment).
GN PSPA.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
OC Streptococcaceae; Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1196;
RA Robinson D.A., Turner J.S., Facklam R.R., Parkinson A.J.,
RA Breiman R.F., Gratten M., Hollingshead S.K., Briles D.E., Crain M.J.;
RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF036620; AAC09239.1; -;
DR EMBL; AF036619; AAC09238.1; -;
FT NON_TER 85
SQ SEQUENCE 85 AA; 9327 MW; CF4C467ACAB4868E CRC64;

Query Match 100.0%; Score 23; DB 2; Length 85;
Best Local Similarity 100.0%; Pred. No. 3.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
||||
DB 45 DYDA 48

RESULT 15
Q97NR9

ID 097NR9 PRELIMINARY; PRT; 86 AA.
AC Q97NR9;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Ribosomal-like PROTEIN-HLA-F product (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93127147; PubMed=1481201;
RA Zemmour J., Parham P.;
RT "A ribosomal protein-like sequence in the 3' untranslated region of
the HLA-F gene";
RL Tissue Antigens 40:250-253(1992).
DR InterPro; IPR001014; Ribosomal_L23.
DR Pfam; PF00276; Ribosomal_L23; 1.
DR ProDom; PD001141; Ribosomal_L23; 1.
DR PROSITE; PS00050; RIBOSOMAL_L23; UNKNOWN_1.
KW MHC.
FT NON_TER 1
FT NON_TER 86
SQ SEQUENCE 86 AA; 9565 MW; 05ED08ED596CD6EE CRC64;

Query Match 100.0%; Score 23; DB 7; Length 86;
Best Local Similarity 100.0%; Pred. No. 3.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
||||
DB 73 DYDA 76

RESULT 16
Q97LK6

ID 097LK6 PRELIMINARY; PRT; 86 AA.
AC 097LK6;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE Hypothetical CF-8 family protein CAC0553.
GN CAC0553.
OS Clostridium acetobutylicum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC Clostridiales; Clostridiaceae; Clostridium.
OX NCBI_TaxID=1488;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
RX MEDLINE=21359325; PubMed=11466286;
RA Noelling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q.,
RA Gibson R., Lee H.M., Dubois J., Qiu D., Hitti J., Wolf Y.I.,
RA Tatuzov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,
RA Bennett G.N., Koonin E.V., Smith D.R.;
RT "Genome sequence and comparative analysis of the solvent-producing
bacterium Clostridium acetobutylicum";
RL J. Bacteriol. 183:4823-4838(2001).
DR EMBL; AE07570; AAC78532.1; -;
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 86 AA; 10155 MW; 004390A546FDA93A CRC64;

Query Match 100.0%; Score 23; DB 16; Length 86;
Best Local Similarity 100.0%; Pred. No. 3.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
||||

```
Db 4 DYDA 7
RESULT 17
Q9RXH2 PRELIMINARY; PRT; 91 AA.
AC Q9RXH2;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Ribosomal protein S15.
GN DR0341.
OS Deinococcus radiodurans.
OC Bacteria; Thermus/Deinococcus group; Deinococci; Deinococcales;
OC Deinococcaceae; Deinococcus.
OX NCBI_TaxID=1299;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=R1;
RX MEDLINE=20036896; PubMed=10567266;
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Vamathevan J.J., Lam P., McDonald L., Utterback T., Zalewski C.,
RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA Fraser C.M.;
RT "Genome sequence of the radioresistant bacterium Deinococcus
RT radiodurans R1.";
RL Science 286:1571-1577(1999).
DR EMBL; AE001894; AAF09921.1; -.
DR HSSP; P80378; IAB3.
DR TIGR; DR0341; -.
DR InterPro; IPR000589; Ribosomal_S15.
DR InterPro; IPR005290; RS15_bact.
DR Pfam; PF00312; Ribosomal_S15; 1.
DR TIGRFAMs; TIGR00952; S15_bact; 1.
DR PROSITE; PS00362; RIBOSOMAL_S15; 1.
KW Ribosomal protein; Complete proteome.
SQ SEQUENCE 91 AA; 10151 MW; AC8601FDF567D2BF CRC64;

Query Match 100.0%; Score 23; DB 16; Length 91;
Best Local Similarity 100.0%; Pred. No. 4.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
DB 70 DYDA 73

RESULT 18
Q72133 PRELIMINARY; PRT; 92 AA.
AC Q72133;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE L1 protein (Fragment).
OS Human papillomavirus.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=10566;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HPVX23;
RX MEDLINE=20295075; PubMed=10834958;
RA Berkhout R.J.M., Bouwes Bavinck J.N., ter Schegget J.;
RT "Persistence of Human Papillomavirus DNA in Benign and (Pre)malignant
RT Skin Lesions from Renal Transplant Recipients.";
RL J. Clin. Microbiol. 38:2087-2096(2000).
DR EMBL; AF054879; AAC12745.1; -.
DR InterPro; IPR002210; PV_capsid_L1.
DR Pfam; PF00500; late_protein_L1; 1.
```

```
DR ProDom; PD000544; PV_capsid_L1; 1.
FT NON_TER 1
SQ SEQUENCE 92 AA; 10490 MW; 9675649BBB8C4901 CRC64;

Query Match 100.0%; Score 23; DB 12; Length 92;
Best Local Similarity 100.0%; Pred. No. 4.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
DB 15 DYDA 18

RESULT 19
Q72134 PRELIMINARY; PRT; 92 AA.
AC Q72134;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE L1 protein (Fragment).
OS Human papillomavirus.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=10566;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HPVX23B;
RX MEDLINE=20295075; PubMed=10834958;
RA Berkhout R.J.M., Bouwes Bavinck J.N., ter Schegget J.;
RT "Persistence of Human Papillomavirus DNA in Benign and (Pre)malignant
RT Skin Lesions from Renal Transplant Recipients.";
RL J. Clin. Microbiol. 38:2087-2096(2000).
DR EMBL; AF054880; AAC12746.1; -.
DR InterPro; IPR002210; PV_capsid_L1.
DR Pfam; PF00500; late_protein_L1; 1.
DR ProDom; PD000544; PV_capsid_L1; 1.
FT NON_TER 1
FT NON_TER 92
SQ SEQUENCE 92 AA; 10531 MW; 37AFFIC96B378931 CRC64;

Query Match 100.0%; Score 23; DB 12; Length 92;
Best Local Similarity 100.0%; Pred. No. 4.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
DB 15 DYDA 18

RESULT 20
Q72138 PRELIMINARY; PRT; 92 AA.
AC Q72138;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE L1 protein (Fragment).
OS Human papillomavirus.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=10566;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HPVX27;
RX MEDLINE=20295075; PubMed=10834958;
RA Berkhout R.J.M., Bouwes Bavinck J.N., ter Schegget J.;
RT "Persistence of Human Papillomavirus DNA in Benign and (Pre)malignant
RT Skin Lesions from Renal Transplant Recipients.";
RL J. Clin. Microbiol. 38:2087-2096(2000).
RN [2]
RP SEQUENCE FROM N.A.
```

```
RC STRAIN=HPVX27;
RA Mulder L.H.C., Berkhout R.J.M., Boxman I.L.A., ter Schegget J.;
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF054884; AAC12750.1; -.
DR InterPro: IPR002210; PV_capsid_L1.
DR Pfam: PF00500; late_protein_L1; 1.
DR ProDom: PD000544; PV_capsid_L1; 1.
FT NON_TER 1
FT NON_TER 92
SQ SEQUENCE 92 AA; 10613 MW; A6F64813AFES0474 CRC64;

Query Match 100.0%; Score 23; DB 12; Length 92;
Best Local Similarity 100.0%; Pred. No. 4.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
DB 15 DYDA 18

RESULT 21
O72141 PRELIMINARY; PRT; 92 AA.
AC O72141;
DT 01-AUG-1998 (TREMBlrel. 07, Created)
DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE L1 protein (Fragment).
OS Human papillomavirus.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=10566;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HPVX33;
RX MEDLINE=20295075; PubMed=10834958;
RA Berkhout R.J.M., Bouwes Bavinck J.N., ter Schegget J.;
RT "Persistence of Human Papillomavirus DNA in Benign and (Pre)malignant
  Skin Lesions from Renal Transplant Recipients.";
RL J. Clin. Microbiol. 38:2087-2096(2000).
DR EMBL: AF054887; AAC12753.1; -.
DR InterPro: IPR002210; PV_capsid_L1.
DR Pfam: PF00500; late_protein_L1; 1.
DR ProDom: PD000544; PV_capsid_L1; 1.
FT NON_TER 1
FT NON_TER 92
SQ SEQUENCE 92 AA; 10633 MW; 06CA4FA913BBD9B8 CRC64;

Query Match 100.0%; Score 23; DB 12; Length 92;
Best Local Similarity 100.0%; Pred. No. 4.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
DB 15 DYDA 18

RESULT 22
O8XE95 PRELIMINARY; PRT; 93 AA.
AC O8XE95;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Hypothetical protein z0510.
GN z0510 OR ECS0462.
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=83334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
```

```
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RN Nature 409:529-533(2001).
RV [2]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / RIMD 0509952;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohemorrhagic Escherichia coli
  O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22(2001).
DR EMBL: AF005220; AAG54758.1; -.
DR EMBL: AF002551; BAB33885.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 93 AA; 10992 MW; 73E7DD97E6153858 CRC64;

Query Match 100.0%; Score 23; DB 16; Length 93;
Best Local Similarity 100.0%; Pred. No. 4.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
DB 51 DYDA 54

RESULT 23
O8ZIU4 PRELIMINARY; PRT; 95 AA.
AC O8ZIU4;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Hypothetical protein YPO0394.
GN YPO0394.
OS Yersinia pestis.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Yersinia.
OX NCBI_TaxID=632;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CO-92 / BIOVAR ORIENTALIS;
RX MEDLINE=21470413; PubMed=11586360;
RA Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,
RA Prentice M.B., Sebaihia M., James K.D., Churcher C., Mungall K.L.,
RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdeno-Tarraga A.M.,
RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
RA Feltwell T., Hamlin N., Holroyd S., Jagels K., Karlyshev A.V.,
RA Leather S., Moule S., Oyston P.C.F., Quail M., Rutherford K.,
RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrall B.G.;
RT "Genome sequence of Yersinia pestis, the causative agent of plague.";
RL Nature 413:523-527(2001).
DR EMBL: AJ414142; CAC89252.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 95 AA; 10766 MW; 234C33AD38A547CC CRC64;

Query Match 100.0%; Score 23; DB 16; Length 95;
Best Local Similarity 100.0%; Pred. No. 4.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
DB 33 DYDA 36
```



```

DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE ZK287.3 protein.
GN ZK287.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA McMurray A.A.;
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology.";
RL Science 282:2012-2018(1998).
DR EMBL; 270757; CAA94799.1; -.
SQ SEQUENCE 102 AA; 11388 MW; 76CBD7C84E6BA695 CRC64;

Query Match 100.0%; Score 23; DB 5; Length 102;
Best Local Similarity 100.0%; Pred. No. 4.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
Db 52 DYDA 55

RESULT 28
O48416
ID O48416 PRELIMINARY; PRT; 103 AA.
AC O48416.
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Protein gp17.
GN 17.
OS Bacteriophage H19B.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Podoviridae.
OX NCBI_TaxID=69932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98343808; PubMed=9680214;
RA Neely M.N.; Friedman D.I.;
RT "Functional and genetic analysis of regulatory regions of coliphage H-
RT 19B: location of Shiga-like toxin and lysis genes suggest a role for
RT phage functions in toxin release.";
RL Mol. Microbiol. 28:1255-1267(1998).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99077833; PubMed=9858702;
RA Neely M.N.; Friedman D.I.;
RT "Arrangement and functional identification of genes in the regulatory
RT region of lambdaoid phage H-19B, a carrier of a Shiga-like toxin.";
RL Gene 223:105-113(1998).
RN [3]
RP SEQUENCE FROM N.A.
RA Neely M.N.; Friedman D.I.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF034975; AAD04642.1; -.
SQ SEQUENCE 103 AA; 11901 MW; 9A76B21E0BACD085 CRC64;

Query Match 100.0%; Score 23; DB 9; Length 103;
Best Local Similarity 100.0%; Pred. No. 4.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
Db 35 DYDA 38

```

```

RESULT 29
O30472
ID O30472 PRELIMINARY; PRT; 104 AA.
AC O30472.
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE YOKX.
GN YOKX.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RA Ghim S.-Y.; Jeong Y.-M.; Choi S.-K.; Park S.-H.;
RT "Sequence analysis of the 30 kb region (182') of the Bacillus subtilis
RT chromosome containing the cge cluster.";
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF006665; AAB81163.1; -.
SQ SEQUENCE 104 AA; 12083 MW; 59293981BC04543A CRC64;

Query Match 100.0%; Score 23; DB 2; Length 104;
Best Local Similarity 100.0%; Pred. No. 4.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
Db 62 DYDA 65

RESULT 30
Q9L9H2
ID Q9L9H2 PRELIMINARY; PRT; 105 AA.
AC Q9L9H2.
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE D-alanine:D-alanine ligase (Fragment).
GN DDL.
OS Enterococcus mundtii.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
OC Enterococcaceae; Enterococcus.
OX NCBI_TaxID=53346;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20383900; PubMed=10930075;
RA Ozawa Y.; Courvalin P.; Galimand M.;
RT "Identification of enterococci at the species level by sequencing of
RT the genes for D-alanine: D-alanine ligases.";
RL Syst. Appl. Microbiol. 23:230-237(2000).
DR EMBL; AF170806; AAF37345.1; -.
DR HSSP; P07862; IIOV.
DR InterPro; IPR000291; Dala_Dala_ligase.
DR Pfam; PF01820; Dala_Dala_ligase; 1.
KW Ligase.
FT NON_TER 1
FT NON_TER 105
SQ SEQUENCE 105 AA; 11725 MW; 8BCA82C22CD8EA2E CRC64;

Query Match 100.0%; Score 23; DB 2; Length 105;
Best Local Similarity 100.0%; Pred. No. 4.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
Db 65 DYDA 68

RESULT 31
O34372
ID O34372 PRELIMINARY; PRT; 105 AA.

```

```

AC O34372:
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DE 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE YODL.
GN YODL.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RA Wambutt R., Wedler H., Lapidus A., Sorokin A., Ehrlich D.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Azevedo V., Berto M.G., Bessieres P., Bolotin A., Borchert S.,
RA Borriess R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Conerton I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,
RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Ghm S.Y., Glaser P., Goffeau A., Gollightly E.J., Grandi G.,
RA Guiseppi G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klaere-Blanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Koningstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA Prescan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadie Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA Sekiguchi J., Sekowska A., Seror S.J., Seror P., Shin B.S., Soldo B.,
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA Viari A., Wambutt R., Wedler E., Wedler H., Wellzenecker T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;
RT "The complete genome sequence of the gram-positive bacterium Bacillus
subtilis.";
RL Nature. 390:249-256(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RA Kunst F., Ogasawara N., Yoshikawa H., Danchin A.;
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF015775; AAB72066.1;
DR EMBL; Z99114; CAB13855.1;
RW Complete proteome.
SQ SEQUENCE 105 AA; 12186 MW; 03BCCA4E80E57663 CRC64;

Query Match 100.0%; Score 23; DB 16; Length 105;
Best Local Similarity 100.0%; Pred. No. 4.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DYDA 4
Db 63 DYDA 66

RESULT 32
Q8YI88
ID Q8YI88 PRELIMINARY; PRT; 107 AA.
AC Q8YI88
DT 01-MAR-2002 (TREMBLrel. 20, Created)

```

```

DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DE 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Hypothetical protein BMEI0557.
GN BMEI0557.
OS Brucella melitensis.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Brucellaceae; Brucella.
OX NCBI_TaxID=29459;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=16M / ATCC 23456 / BIOTYPE 1;
RX MEDLINE=20020109; PubMed=11756688;
RA DelVecchio V.G., Kapratel V., Redkar R.J., Patra G., Mijer C., Los T.,
RA Ivanova N., Anderson I., Bhattacharya A., Lykidis A., Reznik G.,
RA Jablonski L., Larsen N., D'Souza M., Bernal A., Mazur M., Goltsman E.,
RA Selkov E., Elzer P.H., Hagius S., O'Callaghan D., Letesson J.-J.,
RA Haselkorn R., Kypides N., Overbeek R.;
RT "The genome sequence of the facultative intracellular pathogen
Brucella melitensis.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:443-448(2002).
DR EMBL; AE009497; AAL51738.1;
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 107 AA; 11803 MW; 95B1E9462A1EBA03 CRC64;

Query Match 100.0%; Score 23; DB 16; Length 107;
Best Local Similarity 100.0%; Pred. No. 5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DYDA 4
Db 62 DYDA 65

RESULT 33
Q81229
ID Q81229 PRELIMINARY; PRT; 109 AA.
AC Q81229;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Ribosomal protein L25 (Fragment).
OS Zea mays (Maize)
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RX5585; TISSUE=ROOT TIPS;
RA Finkelstein D.B., Drew M.C., Wing R.A., Mullet J.E., Jordan W.R.,
RA Morgan P.W.;
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF061508; AAC24573.1;
DR InterPro; IPR001014; Ribosomal_L23.
DR Pfam; PF00276; Ribosomal_L23; 1.
DR ProDom; PD001141; Ribosomal_L23; 1.
DR PROSITE; PS00050; RIBOSOMAL_L23; 1.
FT NON_TER 1
SQ SEQUENCE 109 AA; 12466 MW; 3F737DD45B5743A2 CRC64;

Query Match 100.0%; Score 23; DB 10; Length 109;
Best Local Similarity 100.0%; Pred. No. 5.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DYDA 4
Db 96 DYDA 99

RESULT 34
Q9A859
ID Q9A859 PRELIMINARY; PRT; 109 AA.
AC Q9A859;

```

DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Transcriptional regulator, ArSR family.
GN CC1505.
OS Caulobacter crescentus.
OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;
OC Caulobacter.
OX NCBI_TaxID=155892;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=ATCC 19089 / CB15;
RC MEDLINE=21173698; PubMed=11259647;
RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R., Ely B.,
RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
RA DeBooy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
RA Kolonay J.F., Smit K., Craven M.B., Khouri H., Shetty J., Berry K.,
RA Utterback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,
RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
RT "Complete genome sequence of Caulobacter crescentus.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
CC -!- SIMILARITY: BELONGS TO THE ARSR FAMILY OF TRANSCRIPTIONAL
CC REGULATORS.
CC EMBL; AE005825; AAK23484.1; -.
DR TIGR; CC1505; -.
DR InterPro; IPR001845; HTH_ARSR.
DR Pfam; PF01022; HTH_5; 1.
DR PRINTS; PRO0778; HTHARSR.
DR SMART; SM00418; HTH_ARSR; 1.
KW DNA-binding; Transcription regulation; Complete proteome.
SQ SEQUENCE 109 AA; 11029 MW; 9A90CFF4AE0A4BC5 CRC64;

Query Match 100.0%; Score 23; DB 16; Length 109;
Best Local Similarity 100.0%; Pred. No. 5.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
Db 68 DYDA 71

RESULT 35
Q9HNJ7
ID Q9HNJ7 PRELIMINARY; PRT; 109 AA.
AC Q9HNJ7
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE Vng2074h.
GN VNG2074h.
OS Halobacterium sp. (strain NRC-1).
OC Archaea; Euryarchaeota; Halobacteriia; Halobacteriales;
OC Halobacteriaceae; Halobacterium.
OX NCBI_TaxID=64091;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20504483; PubMed=11016950;
RA Ng W.V., Kennedy S.P., Mahairas G.G., Berquist B., Pan M.,
RA Shukla H.D., Lasky S.R., Baliga N.S., Thorsson V., Sbrogna J.,
RA Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A.,
RA Leithauser B., Keller K., Cruz R., Danson M.J., Hough D.W.,
RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,
RA Isenbarger T.A., Peck R.F., Pohlschroder M., Spudis J.L., Jung K.-H.,
RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
RA Ehardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.;
RT "Genome sequence of Halobacterium species NRC-1.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
KW EMBL; AE005098; AAG20223.1; -.
SQ SEQUENCE 109 AA; 12102 MW; 27C5648027548E7F CRC64;

Query Match 100.0%; Score 23; DB 17; Length 109;

Best Local Similarity 100.0%; Pred. No. 5.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
Db 18 DYDA 21

RESULT 36
Q8V6T9
ID Q8V6T9 PRELIMINARY; PRT; 110 AA.
AC Q8V6T9
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical 12.6 kDa protein.
OS Halovirus HF2.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae.
OX NCBI_TaxID=33771;
RN [1]
RP SEQUENCE FROM N.A.
RA Tang S.-L., Fisher C., Ngui K., Nuttall S.D., Dyall-Smith M.L.;
RT "Sequence and transcription of halovirus HF2.";
RL Submitted (JAN-2000) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AF222060; AAL54940.1; -.
KW Hypothetical protein.
SQ SEQUENCE 110 AA; 12585 MW; 8B24D378DD0A6BCB CRC64;

Query Match 100.0%; Score 23; DB 12; Length 110;
Best Local Similarity 100.0%; Pred. No. 5.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
Db 22 DYDA 25

RESULT 37
Q9BIF5
ID Q9BIF5 PRELIMINARY; PRT; 111 AA.
AC Q9BIF5
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE Sensory appendage protein-like protein.
GN SAP.
OS Mamestra brassicae.
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC Noctuoidea; Noctuidae; Hadeninae; Mamestra.
OX NCBI_TaxID=55057;
RN [1]
RP SEQUENCE FROM N.A.
RA Jacquelin-Joly E., Francois M.-C., Nagnan-Le Meillour P.;
RT "Molecular cloning of a SAP homolog in Mamestra brassicae male antennae.";
RL Submitted (FEB-2001) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AY026760; AAK14793.1; -.
SQ SEQUENCE 111 AA; 12530 MW; 357877440F9B46AE CRC64;

Query Match 100.0%; Score 23; DB 5; Length 111;
Best Local Similarity 100.0%; Pred. No. 5.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
Db 31 DYDA 34

RESULT 38
O80205
ID O80205 PRELIMINARY; PRT; 112 AA.
AC O80205;

DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last[annotation update])
DE Hypothetical 13.0 kDa protein.
OS Methanobacterium phage psiM2.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae.
OX NCBI_TaxID=77048;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99009353; PubMed=9791169;
RA Pfister P., Wasserfallen A., Stettler R., Leisinger T.;
RT "Molecular analysis of Methanobacterium phage psiM2";
RL Mol. Microbiol. 30:233-244(1998).
RN [2]
RP SEQUENCE FROM N.A.
RX Pfister P., Wasserfallen A., Stettler R., Leisinger T.;
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF065411; AAC27054.1; -;
KW Hypothetical protein.
SQ SEQUENCE 112 AA; 13001 MW; 330A9F4604303994 CRC64;

Query Match 100.0%; Score 23; DB 9; Length 112;
Best Local Similarity 100.0%; Pred. No. 5.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
Db 65 DYDA 68

RESULT 39
Q9CBA7 PRELIMINARY; PRT; 112 AA.
AC Q9CBA7;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE Putative secreted protein.
GN ML2274.
OS Mycobacterium leprae.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1769;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=TN;
RA Cole S.T., Eiglmeyer K., Parkhill J., James K.D., Thomson M.R.,
RA Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D.,
RA Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R.M., Devlin K., Duthoy S., Feltwell T., Fraser A., Hamlin N.,
RA Holroyd S., Hornsby T., Jagels K., Lacroix C., Maclean J., Moule S.,
RA Murphy L., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,
RA Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,
RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
RA Barrell B.G.;
RT "Massive gene decay in the leprosy bacillus.";
RL Nature 409:1007-1011(2001).
DR EMBL; AL583925; CAC31790.1; -;
DR Leproma; ML2274; -;
KW Complete proteome.
SQ SEQUENCE 112 AA; 11815 MW; 133E6D08B4702F5F CRC64;

Query Match 100.0%; Score 23; DB 16; Length 112;
Best Local Similarity 100.0%; Pred. No. 5.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
Db 27 DYDA 30

RESULT 40

O06425 PRELIMINARY; PRT; 112 AA.
AC O06425;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical protein RV0559c.
GN RV0559C OR MTCY25D10.38C OR MT0585.
OS Mycobacterium tuberculosis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=H37RV;
RC MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeyer K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RL Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=H37RV;
RA Parkhill J.;
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / OSHKOSH;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z95558; CAB08970.1; -;
DR EMBL; AE006956; AAK44808.1; -;
DR TIGR; MT0585; -;
DR TubercuList; RV0559c; -;
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 112 AA; 12117 MW; 8A3AF97188684AD4 CRC64;

Query Match 100.0%; Score 23; DB 16; Length 112;
Best Local Similarity 100.0%; Pred. No. 5.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
Db 27 DYDA 30

RESULT 41
Q9HH48 PRELIMINARY; PRT; 113 AA.
AC Q9HH48;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Hypothetical 13.1 kDa protein.
OS Methanobacterium wolfei.
OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;
OC Methanobacteriaceae; Methanothermobacter.
OX NCBI_TaxID=145261;
RN [1]
RP SEQUENCE FROM N.A.

```

RX MEDLINE=21429258; PubMed=11544247;
RT Luo Y., Pfister P., Leisinger T., Wasserfallen A.;
RA "The Genome of Archaeal Prophase Psm100 Encodes the Lytic Enzyme
RT Responsible for Autolysis of Methanothermobacter wolfei.";
RL J. Bacteriol. 183:5788-5792(2001).
RN SEQUENCE FROM N.A.
RP
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF301375; AAC39963.1; -.
KW Hypothetical protein.
SQ SEQUENCE 113 AA; 13102 MW; EF48F6A9FD6C79D2 CRC64;

Query Match 100.0%; Score 23; DB 1; Length 113;
Best Local Similarity 100.0%; Pred. No. 5.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
Db 65 DYDA 68

RESULT 42
QH0N8 PRELIMINARY; PRT; 114 AA.
AC Q9H0N8;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Hypothetical 12.8 kDa protein.
GN DKF2P566F1646.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Carnivora; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN SEQUENCE FROM N.A.
RP
RX TISSUE=KIDNEY;
RC MEDLINE=21154917; PubMed=11230166;
RA Wilmann S., Weil B., Wellenreuther R., Gassenhuber J., Glassl S.,
RA Ansoerge W., Boecker M., Bloecker H., Bauersachs S., Blum H.,
RA Lauber J., Duesterhoeft A., Beyer A., Koehrer K., Strack N.,
RA Mewes H.W., Ottenwaelder B., Obermaier B., Tampe J., Heubner D.,
RA Wambutt R., Korn B., Klein M., Poustka A.;
RT "Towards a Catalog of Human Genes and Proteins: Sequencing and
RT Analysis of 500 Novel Complete Protein Coding Human cDNAs.";
RL Genome Res. 11:422-435(2001).
DR EMBL; ALI36718; CAB66652.1; -.
KW Hypothetical protein.
SQ SEQUENCE 114 AA; 12801 MW; 2D8888AA3551AA86 CRC64;

Query Match 100.0%; Score 23; DB 4; Length 114;
Best Local Similarity 100.0%; Pred. No. 5.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
Db 78 DYDA 81

RESULT 43
Q9AB13 PRELIMINARY; PRT; 114 AA.
AC Q9AB13;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical protein CC0421.
GN CC0421.
OS Caulobacter crescentus.
OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;
OC Caulobacter.
OX NCBI_TaxID=155892;
```

```

RN SEQUENCE FROM N.A.
RP STRAIN=ATCC 19089 / CB15;
RX MEDLINE=21173698; PubMed=11259647;
RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
RA DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,
RA Utterback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,
RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
RT "Complete genome sequence of Caulobacter crescentus.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
DR EMBL; AE005714; AAK22408.1; -.
DR TIGR; CC0421; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 114 AA; 12824 MW; 268FD7B404E4E50D CRC64;

Query Match 100.0%; Score 23; DB 16; Length 114;
Best Local Similarity 100.0%; Pred. No. 5.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
Db 69 DYDA 72

RESULT 44
Q9EXU4 PRELIMINARY; PRT; 117 AA.
AC Q9EXU4;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Nonf-related protein (Fragment).
GN Y20212.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Sinorhizobium.
OX NCBI_TaxID=382;
RN SEQUENCE FROM N.A.
RP STRAIN=RM1021;
RC Finan T.M., Aneja P., Napper K., Golding B.;
RT "Mineral Phosphate Solubilization in Sinorhizobium meliloti.";
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY013584; AAG42546.1; -.
DR InterPro; IPR002818; ThiJ.
DR Pfam; PF01965; ThiJ; 1.
FT NON_TER 117 117
SQ SEQUENCE 117 AA; 12759 MW; CB478AA9C4D9BADD CRC64;

Query Match 100.0%; Score 23; DB 2; Length 117;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
Db 95 DYDA 98

RESULT 45
Q8VHK4 PRELIMINARY; PRT; 117 AA.
AC Q8VHK4;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Aggrecanase-1 (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
```

```
[1]
RN SEQUENCE FROM N.A.
RA Daikou S., Glaeser K.E., Horner A., Davies E.M., Jeffcott L.B.;
RT "Expression of murine aggrecanases.";
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF449483; AAL68774.1;
DR InterPro; IPR001590; Reprolysin.
DR PROSITE; PS50215; ADAM_MEPRO; 1.
FT NON_TER 1
FT NON_TER 117
SQ SEQUENCE 117 AA; 12775 MW; 941853E69F4526C3 CRC64;

Query Match 100.0%; Score 23; DB 11; Length 117;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DYDA 4
    ||||
Db 78 DYDA 81

Search completed: February 6, 2003, 11:21:52
Job time : 24.3333 secs
```

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 6, 2003, 11:16:00 ; Search time 27.8333 Seconds
(without alignments)
19.150 Million cell updates/sec

Title: PAT943-2
Perfect score: 20
Sequence: 1 dlda 4

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues
Total number of hits satisfying chosen parameters: 2536

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 100%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002.*
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	100.0	8	17 AAR97402	Streptococcal M pr
2	20	100.0	8	17 AAR97403	Streptococcal M pr
3	20	100.0	8	22 AAB46882	Integrin alpha-Vbe
4	20	100.0	8	22 AAB46539	Integrin alphav-be
5	20	100.0	8	22 AAB46581	Integrin alphav-be
6	20	100.0	9	17 AAW01152	Mab 1.4 heavy chai
7	20	100.0	9	19 AAW44175	Monoclonal antibod
8	20	100.0	9	22 AAB46868	Integrin alpha-Vbe
9	20	100.0	9	22 AAB46869	Integrin alpha-Vbe
10	20	100.0	9	22 AAB46871	Integrin alpha-Vbe

11	20	100.0	9	22 AAB46872	Integrin alpha-Vbe
12	20	100.0	9	22 AAB46877	Integrin alpha-Vbe
13	20	100.0	9	22 AAB46890	Integrin alpha-Vbe
14	20	100.0	9	22 AAB46891	Integrin alpha-Vbe
15	20	100.0	9	22 AAB46892	Integrin alpha-Vbe
16	20	100.0	10	22 AAG86361	Saccharomyces cere
17	20	100.0	10	22 AAG87624	Saccharomyces cere
18	20	100.0	10	22 AAG87625	Saccharomyces cere
19	20	100.0	10	22 AAB46866	Integrin alpha-Vbe
20	20	100.0	10	22 AAB46867	Integrin alpha-Vbe
21	20	100.0	10	22 AAB46870	Integrin alpha-Vbe
22	20	100.0	10	22 AAB46885	Integrin alpha-Vbe
23	20	100.0	11	22 AAB46862	Integrin alpha-Vbe
24	20	100.0	11	22 AAB46865	Integrin alpha-Vbe
25	20	100.0	11	22 AAB46875	Integrin alpha-Vbe
26	20	100.0	11	22 AAB46878	Integrin alpha-Vbe
27	20	100.0	11	22 AAB46887	Integrin alpha-Vbe
28	20	100.0	12	17 AAR97407	Streptococcal M pr
29	20	100.0	12	17 AAR97408	Streptococcal M pr
30	20	100.0	12	17 AAR97409	Streptococcal M pr
31	20	100.0	12	17 AAR97410	Streptococcal M pr
32	20	100.0	12	22 AAB46888	Integrin alpha-Vbe
33	20	100.0	13	19 AAW72284	Dermatophagoides D
34	20	100.0	13	20 AAY50486	Dermatophagoides s
35	20	100.0	13	20 AAY39913	Paxillin LDI pepti
36	20	100.0	13	22 AAB46889	Integrin alpha-Vbe
37	20	100.0	13	22 AAU19089	-T-cell epitope con
38	20	100.0	14	20 AAY05069	E6 binding domain
39	20	100.0	14	22 AAB46886	Integrin alpha-Vbe
40	20	100.0	15	16 AAR73948	Fragment of T.serg
41	20	100.0	15	17 AAW04908	N. meningitidis 60
42	20	100.0	15	17 AAW04909	N. meningitidis 60
43	20	100.0	15	22 AAG64814	Cytokine-like pept
44	20	100.0	18	22 ABB44423	Peptide #11929 enc
45	20	100.0	18	22 ABB27270	Protein #9269 enco

ALIGNMENTS

RESULT 1
AAR97402
ID AAR97402 standard; peptide; 8 AA.
XX
AC AAR97402;
XX
DT 02-DEC-1996 (first entry)
XX
DE Streptococcal M protein peptide, pl45, fragment 145.1.
XX
DE Streptococcal; M protein; peptide; pl45; chimaeic; chimeric;
KW B-cell; conformational epitope; alpha-helix; GCM4; leucine zipper;
KW detection; mapping; opsonic antibody; vaccine;
KW group A Streptococci; immunotherapy; diagnosis.
XX
OS Streptococcus spp.
XX
PN WO9611944-A1.
XX
PD 25-APR-1996.
XX
PF 16-OCT-1995; 95WO-AU00681.
XX
PR 14-OCT-1994; 94AU-0008851.
XX
PA (BIOT-) BIOTECH AUSTRALIA PTY LTD.
XX (CSIR) COMMONWEALTH SCI & IND RES ORG.
PA (COUN-) COUNCIL QUEENSLAND INST MEDICAL RES.
PA (CSLC-) CSL LTD.
PA (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.
PA (UYME-) UNIV MELBOURNE.
PI Cooper JA, Good ME, Relf WA, Saul AJ;

XX WPI; 1996-221939/22.
XX
XX New chimaeric peptide(s) including a conformational epitope -
PT inserted into a peptide having similar native conformation, useful
PT in vaccines and for determ. of minimal epitope(s) or for mapping
PT amphipathic helices
XX
XX Example 13; Page 37; 99pp; English.
XX
XX The present peptide is a fragment of the Streptococcal M protein
CC peptide p145 (Prusakorn et al, J. Immunol. 149: 2729-2735 (1992)),
CC used in the construction of a novel chimaeric peptide (CP). The CP
CC comprises a B-cell conformational epitope from within p145,
CC inserted into a 2nd peptide, pref. an alpha-helical coil based on
CC the GCN4 leucine zipper peptide (AAR97395). The 2nd peptide has a
CC similar conformation, enabling the epitope to be presented in an
CC immunologically active conformation.
CC The CP can be used in a novel detection/mapping process, e.g. to
CC determine the min. epitope required to induce opsonic antibodies
CC (Ab), and in vaccines against gp. A Streptococci. Ab raised against
CC the CP can be used for immunotherapy and diagnosis, while the CP
CC can be used diagnostically to detect Ab.
XX
XX Sequence 8 AA;
SQ
Query Match 100.0%; Score 20; DB 17; Length 8;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DLDA 4
DB 4 DLDA 7
RESULT 2
AAR97403
ID AAR97403 standard; peptide; 8 AA-
XX
XX AAR97403;
XX
XX 02-DEC-1996 (first entry)
XX
XX Streptococcal M protein peptide, p145, fragment 145.2.
XX
XX Streptococcal; M protein; peptide; p145; chimaeric; chimeric;
KW B-cell; conformational epitope; alpha-helix; GCN4; leucine zipper;
KW detection; mapping; opsonic antibody; vaccine;
KW group A Streptococci; immunotherapy; diagnosis.
XX
XX Streptococcus spp.
XX
XX W09611944-A1.
XX
XX 25-APR-1996.
XX
XX 16-OCT-1995; 95WO-AU00681.
XX
XX 14-OCT-1994; 94AU-0008851.
XX
XX (BIOT-) BIOTECH AUSTRALIA PTY LTD,
PA (CSIR) COMMONWEALTH SCI & IND RES ORG.
PA (COUN-) COUNCIL QUEENSLAND INST MEDICAL RES.
PA (CSLC-) CSL LTD.
PA (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.
PA (UYME) UNIV MELBOURNE.
XX
XX Cooper JA, Good MF, Relf WA, Saul AJ;
XX
XX WPI; 1996-221939/22.
XX
XX New chimaeric peptide(s) including a conformational epitope -
PT inserted into a peptide having similar native conformation, useful

PT in vaccines and for determ. of minimal epitope(s) or for mapping
PT amphipathic helices
XX
XX Example 13; Page 37; 99pp; English.
XX
XX The present peptide is a fragment of the Streptococcal M protein
CC peptide p145 (Prusakorn et al, J. Immunol. 149: 2729-2735 (1992)),
CC used in the construction of a novel chimaeric peptide (CP). The CP
CC comprises a B-cell conformational epitope from within p145,
CC inserted into a 2nd peptide, pref. an alpha-helical coil based on
CC the GCN4 leucine zipper peptide (AAR97395). The 2nd peptide has a
CC similar conformation, enabling the epitope to be presented in an
CC immunologically active conformation.
CC The CP can be used in a novel detection/mapping process, e.g. to
CC determine the min. epitope required to induce opsonic antibodies
CC (Ab), and in vaccines against gp. A Streptococci. Ab raised against
CC the CP can be used for immunotherapy and diagnosis, while the CP
CC can be used diagnostically to detect Ab.
XX
XX Sequence 8 AA;
SQ
Query Match 100.0%; Score 20; DB 17; Length 8;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DLDA 4
DB 2 DLDA 5
RESULT 3
AAB46882
ID AAB46882 standard; peptide; 8 AA.
XX
XX AAB46882;
XX
XX 09-MAY-2001 (first entry)
XX
XX Integrin alpha-Vbeta-6 cyclic peptide inhibitor #14.
XX
XX Cyclic peptide; antithrombotic; cardiant; antiarteriosclerotic;
KW cytotatic; osteopathic; antiinflammatory; antibacterial; antipsoriatic;
KW vulnary; cerebroprotective; antianginal; antidiabetic; nephrotropic;
KW neurothymological; antiarthritic; antirheumatic; antitumor; vasotropic;
KW neuroprotective; alpha_vbeta_6-integrin receptor; thrombosis; tumor;
KW cardiac infarction; coronary heart disease; arteriosclerosis; fibrosis;
KW osteoporosis; inflammation; infection; psoriasis; wound healing.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
FH Misc-difference 5 /note= "D-form residue"
FT
XX
XX DE19933173-A1.
XX
XX 18-JAN-2001.
XX
XX 15-JUL-1999; 99DE-1033173.
XX
XX 15-JUL-1999; 99DE-1033173.
XX
XX (MERE) MERCK PATENT GMBH.
XX
XX Jonczyk A, Diefenbach B, Goodman S;
XX
XX WPI; 2001-192448/20.
XX
XX New cyclic peptides, are alpha(v)beta(6) integrin receptor inhibitors
PT useful e.g. for treating thrombosis, coronary heart disease, tumors,
PT osteoporosis, inflammation or infections
XX
XX Example 1; Page 10; 20pp; German.
PS

XX This invention describes novel cyclic peptides (I), containing 8 amino
 CC acid residues and optionally further alpha,omega-aminocarboxylic acid
 CC residues. Cyclic peptides of formula Cyclo-(Arg-X₁-X₂-X₃-X₄-
 CC X₅-X₆-R₁) (I) and their salts and solvates. X₁ = Ser, Gly or Thr;
 CC X₂ = Leu, Ile, Nle, Val or Phe; X₃ = Asp, Glu, Lys or Phe; X₄ =
 CC Gly, Ala or Ser; X₅ = Leu, Ile, Nle, Val or Phe; X₆ = Arg, His or
 CC Lys; R₁ = one or more omega-aminocarboxylic acid residues, having a
 CC length of 50-2500 mum; or is absent; the amino acid residues are
 CC optionally derivatized and include D- as well as L-forms (in the case of
 CC optically active aminoacids). The products of the invention have
 CC antithrombotic, cardiant, antiarteriosclerotic, cytotatic, osteopathic,
 CC antiinflammatory, antibacterial, antipsoriatic, vulnerary,
 CC cerebroprotective, antianginal, antidiabetic, ophthalmological,
 CC antiarthritic, antirheumatic, antiulcer, vasotropic, nephrotropic,
 CC neuroprotective. (I) are used for treating diseases involving expression
 CC and pathological function of alpha_vbeta_6-integrin receptors, especially
 CC thrombosis, cardiac infarction, coronary heart disease, arteriosclerosis,
 CC tumors, osteoporosis, fibrosis, inflammation, infection, psoriasis and
 CC wound healing deficiency. Other disclosed disorders to be treated include
 CC apoplexy, angina pectoris, ophthalmological diseases (e.g. diabetic
 CC retinopathy, macular degeneration, myopia, ocular histoplasmosis or
 CC rubeotic glaucoma), rheumatoid arthritis, osteoarthritis, ulcerative
 CC colitis, Crohn's disease, atherosclerosis, restenosis after angioplasty,
 CC acute renal failure, renal inflammation or multiple sclerosis. (I) may
 CC also be useful in analytical biology and molecular biology; e.g.
 CC fluorescently labeled (I) may be used as diagnostic markers or (I) may be
 CC used to prepare affinity chromatography columns for purifying integrins.
 CC DNA encoding (I) may be used for treating the same disorders as (I)
 CC itself. (I) are well tolerated and have good alpha_vbeta_6-integrin
 CC receptor inhibiting activity.

XX Sequence 8 AA;

Query Match 100.0%; Score 20; DB 22; Length 8;

Best Local Similarity 100.0%; Pred. No. 7.8e+05;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLDA 4
 ||||
 Db 3 DLDA 6

RESULT 4

AAB46539

ID AAB46539 standard; peptide; 8 AA.

XX AAB46539;

AC AAB46539;

DT 10-APR-2001 (first entry)

Integrin alphav-beta6 inhibitor peptide SEQ ID NO 12.

XX Integrin; inhibitor; thrombolytic; cardiant; antiarteriosclerotic;
 KW cytotatic; osteopathic; antiinflammatory; antimicrobial; antipsoriatic;
 KW vulnerary; antiulcer; cerebroprotective; antianginal; antidiabetic;
 KW ophthalmological; antirheumatic; antiarthritic; vasotropic; nephrotropic;
 KW neuroprotective; thrombosis; coronary infarction; coronary heart disease;
 KW arteriosclerosis; tumor; osteoporosis; fibrosis; inflammation; infection;
 KW psoriasis; wound healing; gastrointestinal disorder; epithelial system;
 KW respiratory disorder.

OS Synthetic.

XX Key Location/Qualifiers

FT Modified-site 1

FT /note= "N-1-Ac-Arg"

FT Modified-site 8

FT /note= "C-terminal amide"

FT

FT

PN DE1929410-A1.

XX

PD 28-DEC-2000.

XX 26-JUN-1999; 99DE-1029410.
 XX 26-JUN-1999; 99DE-1029410.
 PR (MERE) MERCK PATENT GMBH.
 XX Jonczyk A, Diefenbach B, Groth U, Zischinsky G;
 PI WPI; 2001-113366/13.

XX New octapeptide compounds as alpha v beta 6 integrin inhibitors useful
 PT for treating and diagnosing heart disease, tumors, osteoporosis,
 PT fibrosis, inflammation, infection and psoriasis -
 XX Example 1; Page 20; 33pp; German.

XX This invention describes novel octapeptide compounds (I) which have
 CC thrombolytic, cardiant, antiarteriosclerotic, cytotatic, osteopathic,
 CC antiinflammatory, antimicrobial, antipsoriatic, vulnerary, antiulcer,
 CC cerebroprotective, antianginal, antidiabetic, ophthalmological,
 CC antirheumatic, antiarthritic, vasotropic, nephrotropic and
 CC neuroprotective. The compounds were tested for their ability to compete
 CC with fibronectin for binding to alpha_vbeta_6 integrin receptors in
 CC vitro. The compounds are useful as alpha_vbeta_6 integrin inhibitors for
 CC the treatment of thrombosis, coronary infarction, coronary heart disease,
 CC arteriosclerosis, tumors, osteoporosis, fibrosis, inflammation, infection
 CC and psoriasis, and for healing wounds. They can also be used to treat
 CC respiratory and gastrointestinal disorders (e.g. ulcerative colitis or
 CC Crohn's disease), apoplexy, angina pectoris, diabetic retinopathy,
 CC myopia, macular degeneration, ocular histoplasmosis, rheumatoid
 CC arthritis, osteoarthritis, rubeotic glaucoma, post angioplastic
 CC restenosis, acute kidney failure and multiple sclerosis. The compounds
 CC are also useful in diagnosis for detection and localization of
 CC pathological conditions affecting the epithelial system. They can be
 CC conjugated with biotin, radiolabels, marker proteins or antibodies for
 CC this purpose. The compounds are specific and selective for the
 CC alpha_vbeta_6 integrin receptor.

XX Sequence 8 AA;

Query Match 100.0%; Score 20; DB 22; Length 8;

Best Local Similarity 100.0%; Pred. No. 7.8e+05;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLDA 4
 ||||
 Db 3 DLDA 6

RESULT 5

AAB46581

ID AAB46581 standard; peptide; 8 AA.

XX AAB46581;

AC AAB46581;

DT 10-APR-2001 (first entry)

Integrin alphav-beta6 inhibitor peptide #19.

XX Integrin; inhibitor; thrombolytic; cardiant; antiarteriosclerotic;
 KW cytotatic; osteopathic; antiinflammatory; antimicrobial; antipsoriatic;
 KW vulnerary; antiulcer; cerebroprotective; antianginal; antidiabetic;
 KW ophthalmological; antirheumatic; antiarthritic; vasotropic; nephrotropic;
 KW neuroprotective; thrombosis; coronary infarction; coronary heart disease;
 KW arteriosclerosis; tumor; osteoporosis; fibrosis; inflammation; infection;
 KW psoriasis; wound healing; gastrointestinal disorder; epithelial system;
 KW respiratory disorder.

OS Synthetic.

XX Key Location/Qualifiers

FT Modified-site 1

PA (YAMA) YAMANOUCHI PHARM CO LTD.
 XX Hayashi K, Kawauchi Y, Masuho Y, Takasaki J;
 XX WPI; 1998-076914/07.
 XX Amelioration of kidney disorders caused by cisplatin administration
 PT - by treatment with an antibody inhibiting type II phospholipase A2
 PT activity
 XX
 XX Disclosure; Page 41; 74pp; Japanese.
 XX
 XX The present sequence represents a complementary determining region
 CC (CDR) from monoclonal antibody 1.4 heavy chain against type II
 CC phospholipase A2, from the present invention. The present invention
 CC describes a novel method for the amelioration of kidney disorders
 CC (such as acute renal failure) associated with the administration of
 CC cisplatin for the treatment of cancer. The method comprises treatment
 CC with a monoclonal antibody which inhibits the activity of type II
 CC phospholipase A2 (particularly of type II phospholipase A2 of human
 CC origin), or with a protein or peptide possessing the same inhibitory
 CC activity and containing a part of the antibody sequence. Preferably the
 CC antibody also inhibits the activity of ape and/or mouse type II
 CC phospholipase A2, and has the ability to release type II phospholipase
 CC A2 bound to a cell membrane. Three specific monoclonal antibodies
 CC having these properties which can be used are 12H5, 10.1 and 1.4,
 CC derived from hybridomas FERM BP-5300, FERM BP-5298 and FERM BP-5297
 CC respectively. The method can be used for suppressing the nephrotoxicity
 CC which is a characteristic feature of cisplatin administration, and
 CC therefore allowing more efficient use of this drug as an anticancer
 CC agent, e.g. by allowing an increased dosage to be used.
 XX
 XX Sequence 9 AA;
 XX
 XX Query Match 100.0%; Score 20; DB 19; Length 9;
 XX Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 XX Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 DLDA 4
 Db 1 DLDA 4
 IIII
 1 DLDA 4
 RESULT 8
 AAB46868
 ID AAB46868 standard; peptide; 9 AA.
 XX
 XX AAB46868;
 XX
 XX 09-MAY-2001 (first entry)
 XX Integrin alpha-Vbeta-6 cyclic peptide inhibitor #7.
 XX
 XX Cyclic peptide; antithrombotic; cardiant; antiarteriosclerotic;
 KW cytotatic; osteopathic; antiinflammatory; antibacterial; antipsoriatic;
 KW vulnary; cerebroprotective; antiangiinal; antidiabetic; nephrotropic;
 KW ophthalmologic; antiarthritic; antirheumatic; antiulcer; vasotropic;
 KW neuroprotective; alpha_vbeta_6-integrin receptor; thrombosis; tumor;
 KW cardiac infarction; coronary heart disease; arteriosclerosis; fibrosis;
 KW osteoporosis; inflammation; infection; psoriasis; wound healing.
 XX
 XX Synthetic.
 OS
 XX Key Location/Qualifiers
 XX
 XX Misc-difference 5 /note= "D-form residue"
 XX Modified-site 9 /label= Aha
 XX /note= "6-aminohexanoic acid"
 XX
 XX DE19933173-A1.
 PN
 XX 18-JAN-2001.
 PD

XX
 PF 15-JUL-1999; 99DE-1033173.
 XX
 XX 15-JUL-1999; 99DE-1033173.
 PR
 XX (MERE) MERCK PATENT GMBH.
 PA
 XX Jonczyk A, Diefenbach B, Goodman S;
 PI WPI; 2001-192448/20.
 XX
 XX New cyclic peptides, are alpha(v)beta(6) integrin receptor inhibitors
 PT useful e.g. for treating thrombosis, coronary heart disease, tumors,
 PT osteoporosis, inflammation or infections
 XX
 XX Claim 3; Page 5; 20pp; German.
 PS
 XX This invention describes novel cyclic peptides (I), containing 8 amino
 CC acid residues and optionally further alpha-omega-aminocarboxylic acid
 CC residues. Cyclic peptides of formula Cyclo-(Arg-X₁-X₂-X₃-X₄-
 CC X₅-X₆-R₁) (I) and their salts and solvates. X₁ = Ser, Gly or Thr;
 CC X₂ = Leu, Ile, Nle, Val or Phe; X₃ = Asp, Glu, Lys or Phe; X₄ =
 CC Gly, Ala or Ser; X₅ = Leu, Ile, Nle, Val or Phe; X₆ = Arg, His or
 CC Lys; R₁ = one or more omega-aminocarboxylic acid residues, having a
 CC length of 50-2500 mum; or is absent; the amino acid residues are
 CC optionally derivatized and include D- as well as L-forms (in the case of
 CC optically active aminoacids). The products of the invention have
 CC antithrombotic, cardiant, antiarteriosclerotic, cytostatic, osteopathic,
 CC antiinflammatory, antibacterial, antipsoriatic, vulnary,
 CC cerebroprotective, antiangiinal, antidiabetic, ophthalmologic,
 CC antiarthritic, antirheumatic, antiulcer, vasotropic, nephrotropic,
 CC neuroprotective. (I) are used for treating diseases involving expression
 CC and pathological function of alpha_vbeta_6-integrin receptors, especially
 CC thrombosis, cardiac infarction, coronary heart disease, arteriosclerosis,
 CC tumors, osteoporosis, fibrosis, inflammation, infection, psoriasis and
 CC wound healing deficiency. Other disclosed disorders to be treated include
 CC apoplexy, angina pectoris, ophthalmological diseases (e.g. diabetic
 CC retinopathy, macular degeneration, myopia, ocular histoplasmosis or
 CC rubecotic glaucoma), rheumatoid arthritis, osteoarthritis, ulcerative
 CC colitis, Crohn's disease, atherosclerosis, restenosis after angioplasty,
 CC acute renal failure, renal inflammation or multiple sclerosis. (I) may
 CC also be useful in analytical biology and molecular biology; e.g.
 CC fluorescently labeled (I) may be used as diagnostic markers or (I) may be
 CC used to prepare affinity chromatography columns for purifying integrins.
 CC DNA encoding (I) may be used for treating the same disorders as (I)
 CC itself. (I) are well tolerated and have good alpha_vbeta_6-integrin
 CC receptor inhibiting activity.
 XX
 XX Sequence 9 AA;
 XX
 XX Query Match 100.0%; Score 20; DB 22; Length 9;
 XX Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 XX Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 DLDA 4
 Db 3 DLDA 6
 IIII
 3 DLDA 6
 RESULT 9
 AAB46869
 ID AAB46869 standard; peptide; 9 AA.
 XX
 XX AAB46869;
 XX
 XX 09-MAY-2001 (first entry)
 XX Integrin alpha-Vbeta-6 cyclic peptide inhibitor #8.
 XX
 XX Cyclic peptide; antithrombotic; cardiant; antiarteriosclerotic;
 KW cytotatic; osteopathic; antiinflammatory; antibacterial; antipsoriatic;
 KW vulnary; cerebroprotective; antiangiinal; antidiabetic; nephrotropic;
 KW ophthalmologic; antiarthritic; antirheumatic; antiulcer; vasotropic;
 KW osteoporosis; inflammation; infection; psoriasis; wound healing.
 XX
 XX Synthetic.
 OS
 XX Key Location/Qualifiers
 XX
 XX Misc-difference 5 /note= "D-form residue"
 XX Modified-site 9 /label= Aha
 XX /note= "6-aminohexanoic acid"
 XX
 XX DE19933173-A1.
 PN
 XX 18-JAN-2001.
 PD

KW neuroprotective; alpha_vbeta_6-integrin receptor; thrombosis; tumor;
 KW cardiac infarction; coronary heart disease; arteriosclerosis; fibrosis;
 KW osteoporosis; inflammation; infection; psoriasis; wound healing.

XX Synthetic.

XX Key Location/Qualifiers
 FT Misc-difference 5 /note= "D-form residue"
 FT Modified-site 9 /label= Aee
 FT

XX DE19933173-Al.

XX 18-JAN-2001.

XX 15-JUL-1999; 99DE-1033173.

XX 15-JUL-1999; 99DE-1033173.

XX (MERE) MERCK PATENT GMBH.

XX Jonczyk A, Diefenbach B, Goodman S;

XX WPI; 2001-192448/20.

XX New cyclic peptides, are alpha(v)beta(6) integrin receptor inhibitors
 PT useful e.g. for treating thrombosis, coronary heart disease, tumors,
 PT osteoporosis, inflammation or infections

XX Claim 3; Page 5; 20pp; German.

XX This invention describes novel cyclic peptides (I), containing 8 amino
 CC acid residues and optionally further alpha,omega-aminocarboxylic acid
 CC residues. Cyclic peptides of formula Cyclo-(Arg-X₁-Asp-X₁-X₂-X₃-X₄-
 CC X₅-X₆-R₁) (I) and their salts and solvates. X₁ = Ser, Gly or Thr;
 CC X₂ = Leu, Ile, Nle, Val or Phe; X₃ = Asp, Glu, Lys or Phe; X₄ =
 CC Gly, Ala or Ser; X₅ = Leu, Ile, Nle, Val or Phe; X₆ = Arg, Har or
 CC Lys; R₁ = one or more omega-aminocarboxylic acid residues, having a
 CC length of 50-2500 mum; or is absent; the amino acid residues are
 CC optionally derivatized and include D- as well as L-forms (in the case of
 CC optically active aminoacids). The products of the invention have
 CC antithrombotic, cardiant, antiarteriosclerotic, cytostatic, osteopathic,
 CC cerebroprotective, antibacterial, antipsoriatic, vulnerary,
 CC antiarthritic, antirheumatic, antianginal, antidiabetic, ophthalmological,
 CC neuroprotective. (I) are used for treating diseases involving expression
 CC and pathological function of alpha_vbeta_6-integrin receptors, especially
 CC thrombosis, cardiac infarction, coronary heart disease, arteriosclerosis,
 CC tumors, osteoporosis, fibrosis, inflammation, infection, psoriasis and
 CC wound healing deficiency. Other disclosed disorders to be treated include
 CC apoplexy, angina pectoris, ophthalmological diseases (e.g. diabetic
 CC retinopathy, macular degeneration, myopia, ocular histoplasmosis or
 CC rubeotic glaucoma), rheumatoid arthritis, osteoarthritis, ulcerative
 CC colitis, Crohn's disease, atherosclerosis, restenosis after angioplasty,
 CC acute renal failure, renal inflammation or multiple sclerosis. (I) may
 CC also be useful in analytical biology and molecular biology; e.g.
 CC fluorescently labeled (I) may be used as diagnostic markers or (I) may be
 CC used to prepare affinity chromatography columns for purifying integrins.
 CC DNA encoding (I) may be used for treating the same disorders as (I)
 CC itself. (I) are well tolerated and have good alpha_vbeta_6-integrin
 CC receptor inhibiting activity.

XX Sequence 9 AA;
 SQ Query Match 100.0%; Score 20; DB 22; Length 9;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLDA 4

DB 3 DLDA 6

RESULT 10

AAB46871

ID AAB46871 standard; peptide; 9 AA.

XX AAB46871;

XX 09-MAY-2001 (first entry)

XX Integrin alpha-Vbeta-6 cyclic peptide inhibitor #10.

XX Cyclic peptide; antithrombotic; cardiant; antiarteriosclerotic;
 KW cytostatic; osteopathic; antiinflammatory; antibacterial; antipsoriatic;
 KW vulnary; cerebroprotective; antianginal; antidiabetic; nephrotropic;
 KW ophthalmological; antiarthritic; antirheumatic; antiulcer; vasotropic;
 KW neuroprotective; alpha_vbeta_6-integrin receptor; thrombosis; tumor;
 KW cardiac infarction; coronary heart disease; arteriosclerosis; fibrosis;
 KW osteoporosis; inflammation; infection; psoriasis; wound healing.

XX Synthetic.

XX Key Location/Qualifiers
 FT Misc-difference 5 /note= "D-form residue"
 FT Modified-site 9 /note= "beta alanine"
 FT

XX DE19933173-Al.

XX 18-JAN-2001.

XX 15-JUL-1999; 99DE-1033173.

XX 15-JUL-1999; 99DE-1033173.

XX (MERE) MERCK PATENT GMBH.

XX Jonczyk A, Diefenbach B, Goodman S;

XX WPI; 2001-192448/20.

XX New cyclic peptides, are alpha(v)beta(6) integrin receptor inhibitors
 PT useful e.g. for treating thrombosis, coronary heart disease, tumors,
 PT osteoporosis, inflammation or infections

XX Claim 3; Page 5; 20pp; German.

XX This invention describes novel cyclic peptides (I), containing 8 amino
 CC acid residues and optionally further alpha,omega-aminocarboxylic acid
 CC residues. Cyclic peptides of formula Cyclo-(Arg-X₁-Asp-X₁-X₂-X₃-X₄-
 CC X₅-X₆-R₁) (I) and their salts and solvates. X₁ = Ser, Gly or Thr;
 CC X₂ = Leu, Ile, Nle, Val or Phe; X₃ = Asp, Glu, Lys or Phe; X₄ =
 CC Gly, Ala or Ser; X₅ = Leu, Ile, Nle, Val or Phe; X₆ = Arg, Har or
 CC Lys; R₁ = one or more omega-aminocarboxylic acid residues, having a
 CC length of 50-2500 mum; or is absent; the amino acid residues are
 CC optionally derivatized and include D- as well as L-forms (in the case of
 CC optically active aminoacids). The products of the invention have
 CC antithrombotic, cardiant, antiarteriosclerotic, cytostatic, osteopathic,
 CC antiinflammatory, antibacterial, antipsoriatic, vulnerary,
 CC cerebroprotective, antianginal, antidiabetic, ophthalmological,
 CC antiarthritic, antirheumatic, antiulcer, vasotropic, nephrotropic,
 CC neuroprotective. (I) are used for treating diseases involving expression
 CC and pathological function of alpha_vbeta_6-integrin receptors, especially
 CC thrombosis, cardiac infarction, coronary heart disease, arteriosclerosis,
 CC tumors, osteoporosis, fibrosis, inflammation, infection, psoriasis and
 CC wound healing deficiency. Other disclosed disorders to be treated include
 CC apoplexy, angina pectoris, ophthalmological diseases (e.g. diabetic
 CC retinopathy, macular degeneration, myopia, ocular histoplasmosis or
 CC rubeotic glaucoma), rheumatoid arthritis, osteoarthritis, ulcerative
 CC colitis, Crohn's disease, atherosclerosis, restenosis after angioplasty,
 CC acute renal failure, renal inflammation or multiple sclerosis. (I) may
 CC also be useful in analytical biology and molecular biology; e.g.
 CC fluorescently labeled (I) may be used as diagnostic markers or (I) may be

used to prepare affinity chromatography columns for purifying integrins.
DNA encoding (I) may be used for treating the same disorders as (I)
itself. (I) are well tolerated and have good alpha_vbeta_6-integrin
receptor inhibiting activity.

Sequence 9 AA;
Query Match 100.0%; Score 20; DB 22; Length 9;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLDA 4
|
|
|
|
Db 3 DLDA 6

RESULT 11
AAB46872
ID AAB46872 standard; peptide; 9 AA.

AC AAB46872;

DT 09-MAY-2001 (first entry)

DE Integrin alpha-Vbeta-6 cyclic peptide inhibitor #11.

Cyclic peptide; antithrombotic; cardiant; antiarteriosclerotic;
cytostatic; osteopathic; antiinflammatory; antibacterial; antipsoriatic;
vulnerary; cerebroprotective; antiangiinal; antidiabetic; nephrotropic;
ophthalmological; antiarthritic; antirheumatic; antiulcer; vasotropic;
neuroprotective; alpha_vbeta_6-integrin receptor; thrombosis; tumor;
cardiac infarction; coronary heart disease; arteriosclerosis; fibrosis;
osteoporosis; inflammation; infection; psoriasis; wound healing.

OS Synthetic.

Key Location/Qualifiers
FH Misc-difference 5
FT /note= "D-form residue"
FT Modified-site 9
FT /note= "beta alanine"

XX DE19933173-A1.

XX 18-JAN-2001.

XX 15-JUL-1999; 99DE-1033173.

XX 15-JUL-1999; 99DE-1033173.

XX (MERE) MERCK PATENT GMBH.

XX Jonczyk A, Diefenbach B, Goodman S;

XX WPI; 2001-192448/20.

XX New cyclic peptides, are alpha(v)beta(6) integrin receptor inhibitors
PT useful e.g. for treating thrombosis, coronary heart disease, tumors,
PT osteoporosis, inflammation or infections

XX Claim 3; Page 5; 20pp; German.

XX This invention describes novel cyclic peptides (I), containing 8 amino
CC acid residues and optionally further alpha,omega-aminocarboxylic acid
CC residues. Cyclic peptides of formula Cyclo-(Arg-X₁-Asp-X₁-X₂-X₃-X₄-
CC X₅-X₆-R₁) (I) and their salts and solvates. X₁ = Ser, Gly or Thr;
CC X₂ = Leu, Ile, Nle, Val or Phe; X₃ = Asp, Glu, Lys or Phe; X₄ =
CC Gly, Ala or Ser; X₅ = Leu, Ile, Val or Phe; X₆ = Arg, Har or
CC Lys; R₁ = one or more omega-aminocarboxylic acid residues, having a
CC length of 50-2500 mum; or is absent; the amino acid residues are
CC optionally derivatized and include D- as well as L-forms (in the case of
CC optically active aminoacids). The products of the invention have
CC antithrombotic, cardiant, antiarteriosclerotic, cytostatic, osteopathic,

CC antiinflammatory, antibacterial, antipsoriatic, vulnerary,
CC cerebroprotective, antiangiinal, antidiabetic, ophthalmological,
CC antiarthritic, antirheumatic, antiulcer, vasotropic, nephrotropic,
CC neuroprotective. (I) are used for treating diseases involving expression
CC and pathological function of alpha_vbeta_6-integrin receptors, especially
CC thrombosis, cardiac infarction, coronary heart disease, arteriosclerosis,
CC tumors, osteoporosis, fibrosis, inflammation, infection, psoriasis and
CC wound healing deficiency, other disclosed disorders to be treated include
CC apoplexy, angina pectoris, ophthalmological diseases (e.g. diabetic
CC retinopathy, macular degeneration, myopia, ocular histioplasmosis or
CC rubeotic glaucoma), rheumatoid arthritis, osteoarthritis, ulcerative
CC colitis, Crohn's disease, atherosclerosis, restenosis after angioplasty,
CC acute renal failure, renal inflammation or multiple sclerosis. (I) may
CC also be useful in analytical biology and molecular biology; e.g.
CC fluorescently labeled (I) may be used as diagnostic markers or (I) may be
CC used to prepare affinity chromatography columns for purifying integrins.
CC DNA encoding (I) may be used for treating the same disorders as (I)
CC itself. (I) are well tolerated and have good alpha_vbeta_6-integrin
CC receptor inhibiting activity.

XX Sequence 9 AA;

Query Match 100.0%; Score 20; DB 22; Length 9;

Best Local Similarity 100.0%; Pred. No. 7.8e+05;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLDA 4
|
|
|
|
Db 3 DLDA 6

RESULT 12

AAB46877

ID AAB46877 standard; peptide; 9 AA.

XX AAB46877;

XX 09-MAY-2001 (first entry)

DE Integrin alpha-Vbeta-6 cyclic peptide inhibitor SEQ ID 5.

Cyclic peptide; antithrombotic; cardiant; antiarteriosclerotic;
cytostatic; osteopathic; antiinflammatory; antibacterial; antipsoriatic;
vulnerary; cerebroprotective; antiangiinal; antidiabetic; nephrotropic;
ophthalmological; antiarthritic; antirheumatic; antiulcer; vasotropic;
neuroprotective; alpha_vbeta_6-integrin receptor; thrombosis; tumor;
cardiac infarction; coronary heart disease; arteriosclerosis; fibrosis;
osteoporosis; inflammation; infection; psoriasis; wound healing.

OS Synthetic.

XX DE19933173-A1.

XX 18-JAN-2001.

XX 15-JUL-1999; 99DE-1033173.

XX 15-JUL-1999; 99DE-1033173.

XX (MERE) MERCK PATENT GMBH.

XX Jonczyk A, Diefenbach B, Goodman S;

XX WPI; 2001-192448/20.

XX New cyclic peptides, are alpha(v)beta(6) integrin receptor inhibitors
PT useful e.g. for treating thrombosis, coronary heart disease, tumors,
PT osteoporosis, inflammation or infections

XX Examples; Page 16; 20pp; German.

XX This invention describes novel cyclic peptides (I), containing 8 amino
CC acid residues and optionally further alpha,omega-aminocarboxylic acid

CC residues. Cyclic peptides of formula Cyclo-(Arg-X₁-Asp-X₂-X₃-X₄-
 CC X₅-X₆-R₁) (I) and their salts and solvates. X₁ = Ser, Gly or Thr;
 CC X₂ = Leu, Ile, Nle, Val or Phe; X₃ = Asp, Glu, Lys or Phe; X₄ =
 CC Gly, Ala or Ser; X₅ = Leu, Ile, Nle, Val or Phe; X₆ = Arg, His or
 CC Lys; R₁ = one or more omega-aminocarboxylic acid residues, having a
 CC length of 50-2500 mmm; or is absent; the amino acid residues are
 CC optionally derivatized and include D- as well as L-forms (in the case of
 CC optically active amino acids). The products of the invention have
 CC antithrombotic, cardiant, antiarteriosclerotic, cytostatic, osteopathic,
 CC antiinflammatory, antibacterial, antipsoriatic, vulnary, ophthalmological,
 CC cerebroprotective, antiangiogenic, antidiabetic, antihypertensive,
 CC antiarthritic, antirheumatic, antiulcer, vasotropic, nephrotropic,
 CC and neuroprotective. (I) are used for treating diseases involving expression
 CC and pathological function of alpha_vbeta_6-integrin receptors, especially
 CC thrombosis, cardiac infarction, coronary heart disease, arteriosclerosis,
 CC tumors, osteoporosis, fibrosis, inflammation, infection, psoriasis and
 CC wound healing deficiency. Other disclosed disorders to be treated include
 CC apoplexy, angina pectoris, ophthalmological diseases (e.g. diabetic
 CC retinopathy, macular degeneration, myopia, ocular histoplasmosis or
 CC rubecotic glaucoma), rheumatoid arthritis, osteoarthritis, ulcerative
 CC colitis, Crohn's disease, atherosclerosis, restenosis after angioplasty,
 CC acute renal failure, renal inflammation or multiple sclerosis. (I) may
 CC also be useful in analytical biology and molecular biology; e.g.
 CC fluorescently labeled (I) may be used as diagnostic markers or (I) may be
 CC used to prepare affinity chromatography columns for purifying integrins.
 CC DNA encoding (I) may be used for treating the same disorders as (I)
 CC itself. (I) are well tolerated and have good alpha_vbeta_6-integrin
 CC receptor inhibiting activity.

XX Sequence 9 AA;

Query Match 100.0%; Score 20; DB 22; Length 9;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLDA 4
 Db 3 DLDA 6

RESULT 13

AAB46890

ID AAB46890 standard; peptide; 9 AA.

AC AAB46890;

DT 09-MAY-2001 (first entry)

XX Integrin alpha-Vbeta-6 cyclic peptide inhibitor #22.

XX Cyclic peptide; antithrombotic; cardiant; antiarteriosclerotic;

XX cyostatic; osteopathic; antiinflammatory; antibacterial; antipsoriatic;

XX vulnary; cerebroprotective; antiangiogenic; antidiabetic; nephrotropic;

XX ophthalmological; antiarthritic; antirheumatic; antiulcer; vasotropic;

XX neuroprotective; alpha_vbeta_6-integrin receptor; thrombosis; tumor;

XX cardiac infarction; coronary heart disease; arteriosclerosis; fibrosis;

XX osteoporosis; inflammation; infection; psoriasis; wound healing.

XX Synthetic.

XX Key Location/Qualifiers

FT Misc-difference 5

FT /note= "D-form residue"

FT Modified-site 9

FT /label= Aha

FT /note= "6-aminohexanoic acid"

XX DE19933173-A1.

XX 18-JAN-2001.

XX 15-JUL-1999; 99DE-1033173.

XX

PR 15-JUL-1999; 99DE-1033173.

XX (MERE) MERCK PATENT GMBH.

PA Jonczyk A, Diefenbach B, Goodman S;

XX WPI; 2001-192448/20.

DR New cyclic peptides, are alpha(v)beta(6) integrin receptor inhibitors

XX useful e.g. for treating thrombosis, coronary heart disease, tumors,

PT osteoporosis, inflammation or infections

PT Example 1; Page 11; 20pp; German.

XX This invention describes novel cyclic peptides (I), containing 8 amino

XX acid residues and optionally further alpha,omega-aminocarboxylic acid

XX residues. Cyclic peptides of formula Cyclo-(Arg-X₁-Asp-X₂-X₃-X₄-

XX X₅-X₆-R₁) (I) and their salts and solvates. X₁ = Ser, Gly or Thr;

XX X₂ = Leu, Ile, Nle, Val or Phe; X₃ = Asp, Glu, Lys or Phe; X₄ =

XX Gly, Ala or Ser; X₅ = Leu, Ile, Nle, Val or Phe; X₆ = Arg, His or

XX Lys; R₁ = one or more omega-aminocarboxylic acid residues, having a

XX length of 50-2500 mmm; or is absent; the amino acid residues are

XX optionally derivatized and include D- as well as L-forms (in the case of

XX optically active amino acids). The products of the invention have

XX antithrombotic, cardiant, antiarteriosclerotic, cytostatic, osteopathic,

XX antiinflammatory, antibacterial, antipsoriatic, vulnary, ophthalmological,

XX cerebroprotective, antiangiogenic, antidiabetic, antihypertensive,

XX antiarthritic, antirheumatic, antiulcer, vasotropic, nephrotropic,

XX and neuroprotective. (I) are used for treating diseases involving expression

XX and pathological function of alpha_vbeta_6-integrin receptors, especially

XX thrombosis, cardiac infarction, coronary heart disease, arteriosclerosis,

XX tumors, osteoporosis, fibrosis, inflammation, infection, psoriasis and

XX wound healing deficiency. Other disclosed disorders to be treated include

XX apoplexy, angina pectoris, ophthalmological diseases (e.g. diabetic

XX retinopathy, macular degeneration, myopia, ocular histoplasmosis or

XX rubecotic glaucoma), rheumatoid arthritis, osteoarthritis, ulcerative

XX colitis, Crohn's disease, atherosclerosis, restenosis after angioplasty,

XX acute renal failure, renal inflammation or multiple sclerosis. (I) may

XX also be useful in analytical biology and molecular biology; e.g.

XX fluorescently labeled (I) may be used as diagnostic markers or (I) may be

XX used to prepare affinity chromatography columns for purifying integrins.

XX DNA encoding (I) may be used for treating the same disorders as (I)

XX itself. (I) are well tolerated and have good alpha_vbeta_6-integrin

XX receptor inhibiting activity.

XX Sequence 9 AA;

Query Match 100.0%; Score 20; DB 22; Length 9;

Best Local Similarity 100.0%; Pred. No. 7.8e+05;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLDA 4

Db 3 DLDA 6

RESULT 14

AAB46891

ID AAB46891 standard; peptide; 9 AA.

AC AAB46891;

XX 09-MAY-2001 (first entry)

XX Integrin alpha-Vbeta-6 cyclic peptide inhibitor #23.

XX Cyclic peptide; antithrombotic; cardiant; antiarteriosclerotic;

XX cyostatic; osteopathic; antiinflammatory; antibacterial; antipsoriatic;

XX vulnary; cerebroprotective; antiangiogenic; antidiabetic; nephrotropic;

XX ophthalmological; antiarthritic; antirheumatic; antiulcer; vasotropic;

XX neuroprotective; alpha_vbeta_6-integrin receptor; thrombosis; tumor;

XX cardiac infarction; coronary heart disease; arteriosclerosis; fibrosis;

XX osteoporosis; inflammation; infection; psoriasis; wound healing.

```

XX OS Synthetic.
XX ID AAB46892 standard; peptide; 9 AA.
XX AC AAB46892;
XX DT 09-MAY-2001 (first entry)
XX DE Integrin alpha-Vbeta-6 cyclic peptide inhibitor #24.
XX KW Cyclic peptide; antithrombotic; cardiant; antiarteriosclerotic;
XX KW cyostatic; osteopathic; antiinflammatory; antibacterial; antipsoriatic;
XX KW vulnary; cerebroprotective; antianginal; antidiabetic; nephrotropic;
XX KW ophthalmological; antiarthritic; antirheumatic; antitumor; vasotrophic;
XX KW neuroprotective; alpha_vbeta_6-integrin receptor; thrombosis; tumor;
XX KW cardiac infarction; coronary heart disease; arteriosclerosis; fibrosis;
XX KW osteoporosis; inflammation; infection; psoriasis; wound healing.
XX OS Synthetic.
XX FH Key Location/Qualifiers
XX FT Misc-difference 5 /note= "D-form residue"
XX FT Modified-site 9 /label= Abu
XX FT DE19933173-Al.
XX PN 18-JAN-2001.
XX PD 15-JUL-1999; 99DE-1033173.
XX PF 15-JUL-1999; 99DE-1033173.
XX PR 15-JUL-1999; 99DE-1033173.
XX PA (MERE ) MERCK PATENT GMBH.
XX PI Jonczyk A, Diefenbach B, Goodman S;
XX DR WPI; 2001-192448/20.
XX PT New cyclic peptides, are alpha(v)beta(6) integrin receptor inhibitors
XX PT useful e.g. for treating thrombosis, coronary heart disease, tumors,
XX PT osteoporosis, inflammation or infections -
XX PS Example 1; Page 11; 20pp; German.
XX CC This invention describes novel cyclic peptides (I), containing 8 amino
XX CC acid residues and optionally further alpha,omega-aminocarboxylic acid
XX CC residues. Cyclic peptides of formula Cyclo-(Arg-X1-X2-X3-X4-
XX CC X5-X6-R1) (I) and their salts and solvates. X1 = Ser, Gly or Thr;
XX CC X2 = Leu, Ile, Nle, Val or Phe; X3 = Asp, Glu, Lys or Phe; X4 =
XX CC Gly, Ala or Ser; X5 = Leu, Ile, Nle, Val or Phe; X6 = Arg, Har or
XX CC Lys; R1 = one or more omega-aminocarboxylic acid residues, having a
XX CC length of 50-2500 mum; or is absent; the amino acid residues are
XX CC optionally derivatized and include D- as well as L-forms (in the case of
XX CC optically active aminoacids). The products of the invention have
XX CC antithrombotic, cardiant, antiarteriosclerotic, cyostatic, osteopathic,
XX CC antiinflammatory, antibacterial, antipsoriatic, vulnary,
XX CC cerebroprotective, antianginal, antidiabetic, ophthalmological,
XX CC antiarthritic, antirheumatic, antitumor, vasotrophic, nephrotropic,
XX CC neuroprotective. (I) are used for treating diseases involving expression
XX CC and pathological function of alpha_vbeta_6-integrin receptors, especially
XX CC thrombosis, cardiac infarction, coronary heart disease, arteriosclerosis,
XX CC tumors, osteoporosis, fibrosis, inflammation, infection, psoriasis and
XX CC wound healing deficiency. Other disclosed disorders to be treated include
XX CC apoplexy, angina pectoris, ophthalmological diseases (e.g. diabetic
XX CC retinopathy, macular degeneration, myopia, ocular histoplasmosis or
XX CC rubeotic glaucoma), rheumatoid arthritis, osteoarthritis, ulcerative
XX CC colitis, Crohn's disease, atherosclerosis, restenosis after angioplasty,
XX CC acute renal failure, renal inflammation or multiple sclerosis. (I) may
XX CC also be useful in analytical biology and molecular biology; e.g.
XX CC fluorescently labeled (I) may be used as diagnostic markers or (I) may be
XX CC used to prepare affinity chromatography columns for purifying integrins.
XX CC DNA encoding (I) may be used for treating the same disorders as (I)
XX CC itself. (I) are well tolerated and have good alpha_vbeta_6-integrin
XX CC receptor inhibiting activity.
XX SQ Sequence 9 AA;
XX Query Match 100.0%; Score 20; DB 22; Length 9;
XX Best Local Similarity 100.0%; Pred. No. 7.8e+05;
XX Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX QY 1 DLDA 4
XX Db 3 DLDA 6
XX RESULT 15

```

```

AAB46892
ID AAB46892 standard; peptide; 9 AA.
AC AAB46892;
DT 09-MAY-2001 (first entry)
DE Integrin alpha-Vbeta-6 cyclic peptide inhibitor #24.
KW Cyclic peptide; antithrombotic; cardiant; antiarteriosclerotic;
KW cyostatic; osteopathic; antiinflammatory; antibacterial; antipsoriatic;
KW vulnary; cerebroprotective; antianginal; antidiabetic; nephrotropic;
KW ophthalmological; antiarthritic; antirheumatic; antitumor; vasotrophic;
KW neuroprotective; alpha_vbeta_6-integrin receptor; thrombosis; tumor;
KW cardiac infarction; coronary heart disease; arteriosclerosis; fibrosis;
KW osteoporosis; inflammation; infection; psoriasis; wound healing.
OS Synthetic.
FH Key Location/Qualifiers
FT Misc-difference 5 /note= "D-form residue"
FT Modified-site 9 /label= Abu
FT DE19933173-Al.
PN 18-JAN-2001.
PD 15-JUL-1999; 99DE-1033173.
PF 15-JUL-1999; 99DE-1033173.
PR 15-JUL-1999; 99DE-1033173.
PA (MERE ) MERCK PATENT GMBH.
PI Jonczyk A, Diefenbach B, Goodman S;
DR WPI; 2001-192448/20.
PT New cyclic peptides, are alpha(v)beta(6) integrin receptor inhibitors
PT useful e.g. for treating thrombosis, coronary heart disease, tumors,
PT osteoporosis, inflammation or infections -
PS Example 1; Page 11; 20pp; German.
CC This invention describes novel cyclic peptides (I), containing 8 amino
CC acid residues and optionally further alpha,omega-aminocarboxylic acid
CC residues. Cyclic peptides of formula Cyclo-(Arg-X1-X2-X3-X4-
CC X5-X6-R1) (I) and their salts and solvates. X1 = Ser, Gly or Thr;
CC X2 = Leu, Ile, Nle, Val or Phe; X3 = Asp, Glu, Lys or Phe; X4 =
CC Gly, Ala or Ser; X5 = Leu, Ile, Nle, Val or Phe; X6 = Arg, Har or
CC Lys; R1 = one or more omega-aminocarboxylic acid residues, having a
CC length of 50-2500 mum; or is absent; the amino acid residues are
CC optionally derivatized and include D- as well as L-forms (in the case of
CC optically active aminoacids). The products of the invention have
CC antithrombotic, cardiant, antiarteriosclerotic, cyostatic, osteopathic,
CC antiinflammatory, antibacterial, antipsoriatic, vulnary,
CC cerebroprotective, antianginal, antidiabetic, ophthalmological,
CC antiarthritic, antirheumatic, antitumor, vasotrophic, nephrotropic,
CC neuroprotective. (I) are used for treating diseases involving expression
CC and pathological function of alpha_vbeta_6-integrin receptors, especially
CC thrombosis, cardiac infarction, coronary heart disease, arteriosclerosis,
CC tumors, osteoporosis, fibrosis, inflammation, infection, psoriasis and
CC wound healing deficiency. Other disclosed disorders to be treated include
CC apoplexy, angina pectoris, ophthalmological diseases (e.g. diabetic
CC retinopathy, macular degeneration, myopia, ocular histoplasmosis or
CC rubeotic glaucoma), rheumatoid arthritis, osteoarthritis, ulcerative
CC colitis, Crohn's disease, atherosclerosis, restenosis after angioplasty,
CC acute renal failure, renal inflammation or multiple sclerosis. (I) may
CC also be useful in analytical biology and molecular biology; e.g.
CC fluorescently labeled (I) may be used as diagnostic markers or (I) may be
CC used to prepare affinity chromatography columns for purifying integrins.
CC DNA encoding (I) may be used for treating the same disorders as (I)
CC itself. (I) are well tolerated and have good alpha_vbeta_6-integrin
CC receptor inhibiting activity.

```

CC DNA encoding (I) may be used for treating the same disorders as (I)
CC itself. (I) are well tolerated and have good alpha_vbeta_6-integrin
CC receptor inhibiting activity.

XX Sequence 9 AA;

Query Match 100.0%; Score 20; DB 22; Length 9;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLDA 4
Db 3 DLDA 6

RESULT 16

AAAG86361
ID AAG86361 standard; Peptide; 10 AA.

XX AAG86361;

XX 11-SEP-2001 (first entry)

XX Saccharomyces cerevisiae peptide, SEQ ID NO: 1310.

XX Saccharomyces cerevisiae; complementary peptide; peptide identification;
KW drug discovery; drug design.

XX Saccharomyces cerevisiae.

XX WO200142276-A1.

XX 14-JUN-2001.

XX 13-DEC-2000; 2000WO-GB04773.

XX 13-DEC-1999; 99GB-0029471.

XX (PROT-) PROTEOM LTD.

XX Roberts GW, Heal JR;

XX WPI; 2001-367863/38.

XX Identifying complementary peptides by analysis-of protein and
XX nucleotide sequence databases, useful in drug design -

XX Example 3; Page 210; 488pp; English.

XX The invention relates to the identification of complementary peptides
XX by analysis of protein and nucleotide sequence databases from higher
XX eukaryotic genomes, excluding human and plants. The specific
XX complementary peptides interact with their relevant target proteins
XX encoded in the eukaryote genome. The peptides may be used as reagents
XX and drugs for drug discovery and as lead ligands for drug design and
XX development. The present sequence is a complementary peptide from
XX Saccharomyces cerevisiae.

XX Sequence 10 AA;

Query Match 100.0%; Score 20; DB 22; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLDA 4
Db 5 DLDA 8

RESULT 17

AAAG87624
ID AAG87624 standard; Peptide; 10 AA.

XX

AC AAG87624;

XX 11-SEP-2001 (first entry)

XX Saccharomyces cerevisiae peptide, SEQ ID NO: 2573.

XX Saccharomyces cerevisiae; complementary peptide; peptide identification;
KW drug discovery; drug design.

XX Saccharomyces cerevisiae.

XX WO200142276-A1.

XX 14-JUN-2001.

XX 13-DEC-2000; 2000WO-GB04773.

XX 13-DEC-1999; 99GB-0029471.

XX (PROT-) PROTEOM LTD.

XX Roberts GW, Heal JR;

XX WPI; 2001-367863/38.

XX Identifying complementary peptides by analysis of protein and
XX nucleotide sequence databases, useful in drug design -

XX Example 5; Page 384; 488pp; English.

XX The invention relates to the identification of complementary peptides
XX by analysis of protein and nucleotide sequence databases from higher
XX eukaryotic genomes, excluding human and plants. The specific
XX complementary peptides interact with their relevant target proteins
XX encoded in the eukaryote genome. The peptides may be used as reagents
XX and drugs for drug discovery and as lead ligands for drug design and
XX development. The present sequence is a complementary peptide from
XX Saccharomyces cerevisiae.

XX Sequence 10 AA;

Query Match 100.0%; Score 20; DB 22; Length 10;

Best Local Similarity 100.0%; Pred. No. 1.4e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLDA 4
Db 6 DLDA 9

RESULT 18

AAAG87625
ID AAG87625 standard; Peptide; 10 AA.

XX AAG87625;

XX 11-SEP-2001 (first entry)

XX Saccharomyces cerevisiae peptide, SEQ ID NO: 2574.

XX Saccharomyces cerevisiae; complementary peptide; peptide identification;
KW drug discovery; drug design.

XX Saccharomyces cerevisiae.

XX WO200142276-A1.

XX 14-JUN-2001.

XX 13-DEC-2000; 2000WO-GB04773.

XX 13-DEC-1999; 99GB-0029471.

XX